



UNIVERSIDAD DE GRANADA

Programa de
DOCTORADO EN
ESTADÍSTICA MATEMÁTICA Y APLICADA

Modelización e inferencia
sobre procesos de recuento
con intensidad estocástica:
procesos de Cox y auto-excitados

AUTORA:

María del Carmen

MONTES GIJÓN



UNIVERSIDAD DE GRANADA

Modelización e inferencia sobre procesos de recuento
con intensidad estocástica:
procesos de Cox y auto-excitados

AUTORA:
María del Carmen
MONTES GIJÓN

DIRECTORAS:
Dra. Paula
RODRÍGUEZ BOUZAS
y
Dra. Nuria
RUIZ FUENTES

15 de octubre de 2019

Editor: Universidad de Granada. Tesis Doctorales
Autor: María del Carmen Montes Gijón
ISBN: 978-84-1306-443-7
URI: <http://hdl.handle.net/10481/59764>

A Vera

Índice general

Introducción XIII

Objetivos XXV

1. Proceso de Cox 1

1.1. Definición 3

1.2. Estadísticos de recuento 12

1.2.1. Función masa de probabilidad 13

1.2.2. Función característica 14

1.2.3. Funcional característico 16

1.2.4. Moda 19

1.3. Estadísticos de tiempo 20

1.4. Inferencia 21

1.4.1. Inferencia del proceso media 22

1.4.2. Inferencia del proceso intensidad 28

1.4.3. Contraste de bondad de ajuste 31

2. Proceso de Cox compuesto 37

2.1. Definición 38

2.2. Teoremas de representación 40

2.3.	Estadísticos de recuento de un PC compuesto con marcas específicas	43
2.3.1.	Proceso media	43
2.3.2.	Función masa de probabilidad	44
2.3.3.	Moda	45
2.3.4.	Función de distribución acumulada	47
2.4.	Estadísticos de tiempo de un PC compuesto con marcas específicas	47
2.4.1.	Función de supervivencia	47
2.4.2.	Probabilidad de tener una nueva ocurrencia en un intervalo de tiempo	48
2.4.3.	Función de supervivencia condicionada	49
2.4.4.	Densidad de ocurrencia adelantada	49
2.5.	Inferencia	50
2.5.1.	Inferencia de los estadísticos de recuento	50
2.5.2.	Inferencia de los estadísticos de tiempo	52
2.5.3.	Contraste de bondad de ajuste	56
2.6.	Proceso de Cox multidimensional	58
2.6.1.	Procesos de Cox multichannel	58
2.6.2.	Procesos de Cox espacio-tiempo	59
3.	Proceso auto-excitado	61
3.1.	Definición	64
3.1.1.	Intensidad condicionada al recuento	66
3.2.	Estadísticos de recuento	67
3.2.1.	Función masa de probabilidad	67
3.2.2.	Momentos	68
3.3.	Estadísticos de tiempo	70

3.3.1.	Probabilidad de supervivencia	70
3.3.2.	Función de densidad conjunta	71
3.3.3.	Función de densidad muestral	72
3.4.	Estimador de la intensidad condicionada al recuento	74
3.4.1.	Estimación de la función masa de probabilidad	76
3.4.2.	Estimación de la media	77
3.4.3.	Estimación de la intensidad media	79
4.	Simulaciones y Aplicaciones	83
4.1.	Simulaciones. Metodología.	83
4.2.	Simulaciones del PC compuesto	85
4.2.1.	Simulación PC caso 1	86
4.2.2.	Simulación PC caso 2	92
4.2.3.	Simulación PC caso 3	97
4.3.	Aplicaciones del PC compuesto	108
4.3.1.	Aplicación 1: Puntos de cambio de tendencia	108
4.3.2.	Aplicación 2: Datos extremos en metereología	113
4.4.	Simulaciones del proceso auto-excitado	129
4.4.1.	Simulación A-E caso 1	129
4.4.2.	Simulación A-E caso 2	132
A.	Conclusiones	137
A.1.	Proceso de Cox y Cox compuesto	137
A.1.1.	Proceso tipo fase	139
A.2.	Proceso auto-excitado	140
B.	Resumen	143
B.1.	Capítulo 1: Proceso de Cox	145

B.1.1. Contraste de bondad de ajuste dado un PC	147
B.1.2. Contraste de bondad de ajuste dado los datos observa- dos de un PC	148
B.2. Capítulo 2: Proceso de Cox compuesto	149
B.2.1. Contraste de bondad de ajuste de un PC compuesto con marcas específicas	151
B.3. Capítulo 3: Procesos auto-excitados	153
B.4. Capítulo 4: Simulaciones y aplicaciones	154
Bibliografía	157

Índice de figuras

1.1. Trayectoria muestral de la intensidad como un proceso tipo fase.	12
2.1. Proceso de Cox compuesto con marcas asociadas.	38
2.2. Proceso de Cox compuesto con marcas en un subconjunto B.	39
3.1. Diagrama de las trayectorias muestrales del proceso y sus puntos estimada la intensidad de recuento condicionada en un instante de tiempo.	75
3.2. Diagrama del estimador de la intensidad de recuento condicio- nada dadas $n-1$ y n ocurrencias cuando los intervalos donde las funciones son distintas de cero se superponen o no.	79
3.3. Diagrama para ilustrar los tiempos ω_S y ω_J	80
4.1. Simulación PC 1. Gráfico de las estimaciones de la probabilidad de tener una nueva ocurrencia antes de un tiempo s para $T_1 = 10$ del proceso $N(t, B)$	92
4.2. Simulación PC 2. Gráfico de las estimaciones de la probabilidad de tener una nueva ocurrencia antes de un tiempo s para $T_1 = 10$ del proceso $N(t, B)$	97

4.3. Simulación PC 3. Gráfico del espacio de marcas con su distribución y el conjunto de marcas B del proceso $N(t, B)$	99
4.4. Simulación PC 3. Trayectorias muestrales del PC sin tener en cuenta las marcas y el PC con marcas en el conjunto B del proceso $N(t, B)$	99
4.5. Simulación PC 3. Gráfico de las estimaciones de la probabilidad de tener una nueva ocurrencia antes de un tiempo s para $T_1 = 10$ del proceso $N(t, B)$	104
4.6. Simulación PC 3. 100 nuevas trayectorias de la media del proceso intensidad.	105
4.7. Simulación PC 3. 3 de las nuevas trayectorias de la media del proceso intensidad.	106
4.8. Aplicación 1. Gráfico de cajas y bigotes del RECM de la estimación de la media con todas las trayectorias muestrales para distintos valores de T_1	110
4.9. Aplicación 1. Trayectorias muestrales de los cambios de tendencia de las acciones ALPHA, CAT, EK, etc.	113
4.10. Aplicación 1. Gráfico de cajas y bigotes del RECM de la estimación de la media entre todas las trayectorias muestrales para distintos valores de T_1	114
4.11. Aplicación 2. Provincias y zonas climáticas de España.	115
4.12. Aplicación 2. Temperaturas máximas diarias, extremos máximos y extremos máximos en B en la estación meteorológica de Sevilla en 2011.	117
4.13. Aplicación 2. Temperaturas máximas diarias, extremos máximos y extremos en B de la estación meteorológica de Sevilla durante los primeros 80 días del 2011.	117

4.14. Aplicación 2. Trayectorias muestrales del proceso de recuento de los extremos de temperaturas máximas y los extremos en el subconjunto B de todas las estaciones meteorológicas.	118
4.15. Aplicación 2. Temperaturas mínimas diarias, extremos mínimos y extremos en B de la provincia de Sevilla en 2011.	121
4.16. Aplicación 2. Temperaturas mínimas diarias, extremos mínimos y extremos en B de la provincia de Sevilla durante los primeros 80 días del 2011.	122
4.17. Aplicación 2. Trayectorias muestrales del proceso de recuento de los extremos de temperatura mínima y los extremos en el subconjunto B de todas las estaciones meteorológicas.	122
4.18. Simulación A-E 1. Tiempos de ocurrencias de las tres primeras ocurrencias.	130
4.19. Simulación A-E 1. Estimación intensidad condicionada al recuento.	130
4.20. Simulación A-E 1. Estimación intensidad condicionada al recuento $n=0$	131
4.21. Simulación A-E 1. Estimación intensidad condicionada al recuento $n=9$	132
4.22. Simulación A-E 2. Trayectorias muestrales de proceso auto-excitado con intensidad tipo fase.	134
4.23. Simulación A-E 2. Tiempos de ocurrencias de las tres primeras ocurrencias.	135
4.24. Simulación A-E 2. Estimación intensidad condicionada al recuento.	135
4.25. Simulación A-E 2. Estimación intensidad condicionada al recuento $n=1$	136
4.26. Simulación A-E 2. Estimación intensidad condicionada al recuento $n=9$	136

Índice de tablas

4.1. Simulación PC 1. Estimación de la media, moda y probabilidad de tener una nueva ocurrencia antes de un tiempo s (Prob) del proceso $N(t, B)$	89
4.2. Simulación PC 1. Comparativa de las medias estimadas y las medias muestrales para cada trayectoria del proceso $N(t, B)$	90
4.3. Simulación PC 1. Valores más frecuentes del número de puntos antes de s y sus frecuencias del proceso $N(t, B)$ para $T_1 = 10$	91
4.4. Simulación PC 2. Estimación de la media, moda y probabilidad de tener una nueva ocurrencia antes de un tiempo s (Prob) del proceso $N(t, B)$	94
4.5. Simulación PC 2. Comparativa de las medias estimadas con las medias muestrales del proceso $N(t, B)$ para cada trayectoria de $N(s, B)$	95
4.6. Simulación PC 2. Valores más frecuentes del número de puntos antes de s y sus frecuencias del proceso $N(t, B)$ para $T_1 = 10$	96
4.7. Simulación PC 3. Estimación de la media, moda y probabilidad de tener una nueva ocurrencia antes de un tiempo s (Prob) del proceso $N(t, B)$	101

4.8. Simulación PC 3. Comparativa de las medias estimadas y las medias muestrales para cada trayectoria de $N(s, B)$ del proceso $N(t, B)$.	102
4.9. Simulación PC 3. Valores más frecuentes del número de puntos antes de s y sus frecuencias del proceso $N(t, B)$ para $T_1 = 10$.	103
4.10. Simulación PC 3. Porcentaje de trayectorias no rechazadas por el contraste de bondad de ajuste sin corromper o eliminando cada vez más puntos.	106
4.11. Simulación PC 3. Porcentaje de trayectorias no rechazadas por cada contraste de bondad de ajuste según el valor de la probabilidad p .	107
4.12. Aplicación 1. Media y moda para diferentes valores de s para $T_1 = 15$, error y ajuste.	110
4.13. Aplicación 1. Media y moda para diferentes valores de s y T_1 , error y ajuste.	111
4.14. Aplicación 1. Media y desviación típica del RECM (ε) de las estimaciones medias del número de cambios de tendencia a partir de todas las trayectorias muestrales.	112
4.15. Aplicación 1. Media y desviación típica del RECM (ε) de las estimaciones de la media para el número de cambios de tendencia entre todas las trayectorias muestrales.	114
4.16. Aplicación 2. Estimación de la media, moda y probabilidad de tener un nuevo punto antes de un tiempo s para los valores extremos de temperatura máxima en el subconjunto B .	120
4.17. Aplicación 2. Comparativa de las medias estimadas y las medias muestrales de $N(s, B)$ para los extremos de temperatura máxima en B .	121

-
- 4.18. Aplicación 2. Estimación de la media, moda y probabilidad de tener un nuevo punto antes de s para los extremos de temperaturas mínimas en el subconjunto B 124
- 4.19. Aplicación 2. Comparativa de la estimación de la media y medias muestrales de $N(s, B)$ para las temperaturas mínimas. 125
- 4.20. Aplicación 2. P-valor para el test de bondad de ajuste sobre las temperaturas mínimas. 125
- 4.21. Aplicación 2. Evaluación de si la nueva trayectoria pertenece al proceso original de temperaturas mínimas para Zaragoza y Orense. 126
- 4.22. Aplicación 2. P-valor para el test de bondad de ajuste sobre las temperaturas máximas. 127
- 4.23. Aplicación 2. Evaluación de si la nueva trayectoria pertenece al proceso original de temperaturas máximas para Zaragoza y Orense. 128
- 4.24. Simulación A-E 1. Estimación de la media del proceso y de la media de la estimación de la intensidad. 133
- 4.25. Simulación A-E 1. Estimación de la función masa de probabilidad. 134

Introducción

Un fenómeno en el que un suceso ocurre repetidas veces en el tiempo o el espacio se modeliza matemáticamente, atendiendo a las ocurrencias, según un proceso puntual. Sin embargo, pueden tomarse distintas perspectivas en el estudio de este tipo de fenómenos. Al margen de que el ser humano siempre ha estado interesado en observar ocurrencias y contarlas, como catástrofes naturales (inundaciones, terremotos,...), estrellas, etc., los orígenes de la actual teoría de procesos puntuales no son fáciles de determinar.

De todas las piezas que hoy conforman la teoría de los procesos puntuales, la más antigua es la asociada al estudio de los intervalos entre ocurrencias, es decir, la teoría de procesos de renovación. El estudio del tiempo entre fallos de un sistema se desarrolló durante los años 30 del siglo XX y alcanzó su estado de madurez gracias a William Feller (1906-1970) y otros durante la Segunda Guerra Mundial. Las tres áreas de aplicación más significativas son la teoría de colas, la teoría actuarial y la teoría de fiabilidad.

El otro acercamiento básico y clásico a los procesos puntuales es el recuento del número de sucesos, puntos u ocurrencias, en intervalos o regiones de un espacio dando lugar a la denominación de proceso de recuento. En este ámbito, la teoría de probabilidad de distribuciones discretas juega un papel primordial. El punto de partida de este punto de vista de un proceso puntual

es el texto que Siméon Denis Poisson (1781–1840) publicó en 1837 [97], el hallazgo de la distribución de Poisson como el límite de una binominal. Las primeras discusiones de problemas de recuento se deben a Seidel, 1876 [110] y a Abbe, 1879 [2], que trataron las ocurrencias de tormentas y el número de células sanguíneas en áreas cuadradas de un hemocitómetro, respectivamente, y parece que de forma independiente con respecto al trabajo de Poisson. Más tarde, Erlang, 1909 [48] obtuvo la distribución de Poisson para el número de llamadas entrantes en una centralita telefónica. En 1910, Bateman [11] utilizó las matemáticas para expresar el problema de Rutherford y Geiger de recuento de partículas como una familia de ecuaciones diferenciales cuya solución eran las probabilidades de una distribución de Poisson. Estas ecuaciones representan un proceso de nacimiento puro en la teoría de procesos de nacimiento y muerte, teoría que se desarrolló notablemente en las dos siguientes décadas. Igualmente ocurrió con los procesos de ramificación. La conexión entre los procesos de nacimiento y muerte y los de ramificación con los procesos puntuales o de recuento comenzó a ser importante a partir de los años 40. El proceso de recuento más básico sería el proceso de Poisson y sigue considerándose la base sobre la que construir otros más complejos.

En muchos problemas de ecología, astronomía, etc., a menudo se observa que la dispersión de los puntos es mayor que en una distribución de Poisson. Estudiando estadísticas de accidentes, Greenwood, 1920 [61] obtuvo un modelo al que denominaron distribución de Poisson mixta en la que el parámetro de la distribución de Poisson se consideraba una variable aleatoria.

Durante la Segunda Guerra Mundial y sobre todo después, se dió un desarrollo espectacular en la teoría y aplicaciones de los procesos estocásticos y, en concreto, de los que nos ocupan. Se promovió un interés creciente por aplicar las matemáticas más rigurosas a procesos puntuales y su utilización en

aplicaciones. Moyal, 1962 [88], en el contexto de problemas físicos y ecológicos, fue el que por primera vez construyó la matemática necesaria para los procesos puntuales en un espacio de estados general, clarificando las relaciones entre distribuciones finito-dimensionales y funcionales característicos y de probabilidad.

Paralelamente, el artículo de Cox, 1955 [40] contenía el origen de muchas cuestiones relevantes para los procesos puntuales y se discutían varios modelos diferentes incluyendo el importante proceso de Poisson doblemente estocástico, hoy más conocido como proceso de Cox tomando su nombre. Este nuevo proceso englobaba a todos los anteriores como casos particulares. Su novedad clave era la naturaleza estocástica de su intensidad ya que era a su vez un proceso (de ahí la denominación de doblemente estocástico). Este hecho confiere al proceso puntual un importante grado de flexibilidad. A partir de este trabajo y especialmente en los años 70, se suscitó un creciente interés por el uso de procesos puntuales en ingeniería de comunicaciones y medicina (detección de señales, impulsos nerviosos, reconstrucción de imágenes, etc.). En particular, se empieza a focalizar ese interés en la modelización y predicción de estos procesos.

En el último cuarto del siglo XX se han desarrollado o consolidado campos como análisis de supervivencia, procesos puntuales espaciales, procesos puntuales en el espacio-tiempo, y procesos puntuales marcados. Es más, la conexión entre procesos puntuales e historiales de sucesos, observados como sucesos recurrentes o bien con la estructura de datos de recuento de panel, se está estrechando cada vez más. La introducción de Aalen et al., 2008 [1] explica magistralmente la relación entre procesos puntuales, de recuento, análisis de supervivencia y análisis de historiales de sucesos.

El recuento de terremotos estudiado por Ogata, 1998 [92] o Gospodinov y Pavlova, 2001 [58] es el ejemplo por excelencia con el que se suele ilustrar la aplicación del proceso de Cox a datos reales. Desde entonces, muchos autores han estudiado el proceso de Cox desde distintos puntos de vista como se verá poco a poco a lo largo de esta memoria. Sin embargo hay que destacar Snyder y Miller, 1991 [117] como texto fundamental y completo para estudiar la evolución del proceso de Poisson hasta el proceso de Cox pasando por el proceso de Poisson compuesto, trasladado, filtrado, etc.

La estimación del proceso de Cox la han estudiado también muchos otros autores, pero habitualmente asumiendo una cierta estructura de su intensidad. Para el propósito de esta memoria, interesa especialmente el texto de Bouzas, 2011 [28]. En él, usando la perspectiva de Snyder y Miller, se analizan generalizaciones del proceso de Cox, pero lo más novedoso es el uso de la potente técnica de análisis funcional en componentes principales para realizar estimaciones y predicciones específicamente pensadas para este proceso. Más adelante, Bouzas y Ruiz-Fuentes, 2015 [31] hace un repaso de la inferencia llevada a cabo con esta técnica hasta ese momento tanto para el proceso de Cox como para el proceso de Cox compuesto. Lo interesante de esta línea de inferencia, es el hecho de poder realizarse a partir de datos observados del proceso de recuento sin presunciones sobre la estructura de su proceso intensidad, lo cual es habitualmente más realista que la imposición de un modelo prefijado.

Este repaso al origen y evolución de los procesos puntuales o de recuento aporta varias ideas generales interesantes. La más simple es el hecho de que realmente, la teoría de procesos de recuento no despegó como tal hasta los años 70 del siglo pasado, es decir, hace unos 50 años aproximadamente. Es también importante subrayar la necesidad continua de generalización de los

procesos puntuales para acoger la mayor cantidad posible de casos reales y modelizarlos lo más fielmente que sea factible. En este sentido, es crucial la aportación de la doble aleatoriedad que incluye el proceso de Cox. Y por último, como es evidente y tal como ocurre en cualquier campo dentro de la estadística y la probabilidad, es imprescindible hacer inferencia sobre estos procesos.

Después de lo expuesto, parece claro el interés de los procesos de recuento doblemente estocásticos. Esta memoria de doctorado se adentra en el estudio de este tipo de procesos comenzando por el proceso de Cox. A pesar de que este proceso está ya bastante estudiado, el hecho de que las expresiones de sus estadísticos sean muy complicadas excepto para algunos casos concretos, aún plantea muchas incógnitas y áreas de estudio. Sin embargo, alguna de sus propiedades pueden resultar restrictivas para conjuntos de datos reales complejos. Por ello, otro propósito más es generalizar el proceso de Cox relajando alguna de esas propiedades. La primera, mediante la inclusión de marcas para cada suceso, dando lugar al proceso de Cox compuesto, y la segunda, añadiendo la posibilidad de dependencia del pasado, surgiendo así los procesos auto-excitados.

Más concretamente, en el Capítulo 1, Proceso de Cox, se define este proceso doblemente estocástico según la perspectiva clásica reflejada en Snyder y Miller, 1991 [117]. Según esta definición, el proceso queda caracterizado mediante el proceso intensidad. Sin embargo, se recuerda también su posible definición mediante el proceso media que tendrá un papel fundamental tanto para su estimación como su predicción. Esto significa que proponiendo distintos procesos intensidad o media, se construyen diferentes casos de un proceso de Cox. En aras de la búsqueda de generalizaciones que se ha comentado, se propone un nuevo caso para el proceso intensidad, el que se

ha denominado como proceso tipo fase. Éste será un proceso basado en una distribución tipo fase cuyos parámetros son a su vez aleatorios. El proceso de Cox con esta intensidad resulta especialmente flexible y por tanto interesante para la modelización de procesos de recuento reales. Esta contribución ha sido publicada en Bouzas et al., 2019 [33].

A continuación se presentan los estadísticos de recuento y de tiempo de un proceso de Cox. Puesto que todos ellos están caracterizados por el proceso intensidad o el proceso media, es clara la importancia de la estimación y predicción de cualquiera de ellos para hacer inferencia sobre los primeros. Este es el primer motivo por el que se repasa la inferencia de ambos, intensidad y media, usando análisis de componentes principales. Una segunda razón es que, a su vez, esto nos permite construir un contraste de bondad de ajuste para un proceso de Cox.

La pregunta clásica que surge continuamente en estadística de si unos datos siguen o no una determinada distribución se traduce en este marco de trabajo de la siguiente manera. Habiendo observado una trayectoria de un proceso de recuento, ¿se ajusta a un cierto proceso de Cox propuesto? Bouzas et al., 2010b [26] ya formuló un contraste inspirado en el problema de calibración de un contador de partículas emitidas por isótopos radiactivos. En este caso, los datos estudiados son datos reales de los que se desconoce su estructura estocástica. En este capítulo de la memoria, se aporta un contraste análogo pero para aquellos casos de proceso de Cox con proceso intensidad conocido. Este nuevo caso de contraste de bondad de ajuste se ha presentado en Ruiz-Fuentes et al. ERCIM 2016, [106].

En el Capítulo 2, Proceso de Cox compuesto, se estudia la generalización del proceso de Cox añadiendo marcas a sus ocurrencias. Cada marca es una variable aleatoria que puede tener diferentes significados como el lugar en

el espacio en el que se da esa ocurrencia, la procedencia de la misma, su tamaño, etc. Claramente, la casuística es interminable, por lo que estamos antes un proceso que puede modelizar una gran cantidad de realidades distintas y peculiares. Una interpretación especialmente interesante es el número de puntos que están ocurriendo en ese instante. De esta forma, se permite la ocurrencia simultánea de sucesos o puntos y por lo tanto se relaja la propiedad de regularidad. Como el título del capítulo indica, este es el conocido como proceso de Cox compuesto.

Suele ser habitual que, por diferentes causas, interese estudiar algunas de las ocurrencias en lugar de todas ellas. En otras palabras, puede que interese estudiar solo las ocurrencias que comparten un cierto tipo de marca específica (pertenecen a una zona, tienen un cierto tamaño, etc.). En esta memoria, gracias a los teoremas de representación, se recogen los estadísticos de recuento para estos procesos y su estimación. Asimismo, se aportan expresiones cerradas de sus estadísticos de tiempo y sus correspondientes estimaciones que han sido publicadas en Bouzas et al., 2019 [33].

Por último, los contrastes de bondad de ajuste para el PC se han extendido al PC compuesto tanto para el caso de proceso intensidad conocido como desconocido. Estos tests se han presentado en Ruiz Fuentes et al., ERCIM 2016, [106].

El Capítulo 3, Proceso auto-excitado, pretende generalizar el proceso de Cox desde otra perspectiva, introduciendo la posibilidad de que el proceso dependa de su pasado. Es decir, cada nueva ocurrencia, podría depender del número de ocurrencias anteriores, de los tiempos de ocurrencia de alguna o todas ellas, etc. No es nuevo el estudio de este proceso, véase por ejemplo el proceso de Hawkes como modelo relacionado y en cierto modo predecesor del auto-excitado más general. Sin embargo, es curioso cómo hoy en día con-

curren las dos siguientes circunstancias. A pesar de que están aumentando notablemente las publicaciones sobre procesos autoexcitados, suelen restringirse a casos cuya intensidad sea combinación lineal de variables aleatorias o similares y, por otra parte, suelen ser procesos espaciales. Lo primero, sin duda, es una restricción muy fuerte, y lo segundo, *a priori* puede parecer general pero no tiene por qué serlo. Siguiendo la línea de los anteriores capítulos, en esta memoria no se presupone una estructura o distribución prefijada para la intensidad, sino que se estima a partir de los datos observados del proceso de recuento.

Partiendo nuevamente de la definición de proceso auto-excitado de Snyder y Miller, 1991 [117] y gracias a una sugerencia en ese libro, se proporcionan expresiones explícitas de varios estadísticos básicos de recuento y se repasan los de tiempo. Surge de forma natural un especial interés por la intensidad condicionada al recuento por ser un proceso que caracteriza a los estadísticos. Como consecuencia, la estimación de esta intensidad condicionada al recuento resulta no solo interesante en sí misma, sino fundamental para hacer inferencia sobre los estadísticos del proceso auto-excitado. En esta memoria, se construye un estimador puntual para la intensidad condicionada al recuento y, a partir de ella, se obtienen estimaciones para la función masa de probabilidad y la media de proceso, así como para la media del proceso intensidad.

Con respecto al Capítulo 4, Simulaciones y Aplicaciones, ha de hacerse alguna salvedad. Es imprescindible ilustrar cada aportación de la memoria con ejemplos de simulación, aplicaciones o ambos y por supuesto probar el buen funcionamiento de los contrastes propuestos. Dado que para un mismo ejemplo de proceso de recuento se pueden aplicar muchas de las distintas contribuciones, a fin de recoger la aplicación de varias de ellas para cada uno

de los ejemplos de procesos escogidos y presentarlo como un todo, se han incluido todas las simulaciones y aplicaciones en este capítulo.

En una primera sección se simulan varios casos de proceso de Cox compuesto y se muestra la estimación del proceso media, la moda y la probabilidad de tener un nuevo suceso en un intervalo prefijado. La precisión de todas las estimaciones se evalúa para probar su fiabilidad obteniéndose errores de estimación muy pequeños. Los ejemplos están escogidos cuidadosamente para ilustrar que el proceso de Cox compuesto puede modelizar muy diversos tipos de procesos de recuento como el proceso de Cox con supresiones, el formado por procesos de Poisson simultáneos o proceso de Cox tiempo-espacio. En el último caso se añade también la particularidad de que el proceso intensidad es un proceso tipo fase.

Debe apuntarse también que no se simula un proceso de Cox puesto que este es caso particular del de Cox y, por otra parte, como se verá más adelante un proceso de Cox compuesto con marcas específicas en un subconjunto del espacio de marcas es un proceso de Cox. Por tanto, se evita la redundancia en cuanto a simulaciones y su uso en la aplicación de los resultados.

El contraste de bondad de ajuste para un proceso de Cox compuesto propuesto en el Capítulo 2, generalización del propuesto para un proceso de Cox en el Capítulo 1 es evaluado para demostrar su eficacia. En este punto de la memoria, se simulan varios ejemplos de procesos y se les contamina de diversas maneras antes de aplicarles el contraste. Se comprueba que en un alto porcentaje de veces el test detecta cuándo la trayectoria observada no sigue el modelo propuesto y, por tanto, queda probada la eficacia del test.

Los resultados teóricos de esta memoria han sido utilizados para responder a cuestiones planteadas por conjuntos de datos reales. No es fácil encontrar datos adecuados o trabajar con ellos, pero se han podido estudiar datos

de puntos de cambio de tendencia en cotizaciones bursátiles y datos extremos en temperaturas. Estos ejemplos dan una idea de posibles aplicaciones del trabajo realizado en esta memoria.

Finalmente, se simulan procesos auto-excitados con varias intensidades incluyendo el caso del proceso tipo fase. Aplicando los resultados del Capítulo 3, se estiman su intensidad condicionada al recuento, la función masa de probabilidad, la media y la media de la intensidad. Es de destacar que la mayor dificultad en el cálculo de estas estimaciones es la resolución de integrales dobles que se encadenan. Se ha utilizado integración numérica para resolver esos problemas de cálculo.

A la luz de este último capítulo, se puede concluir que se ha dado un paso más en el estudio de procesos de recuento doblemente estocásticos obteniendo resultados teóricos para la estimación de estadísticos de estos procesos y se han propuesto contrastes de bondad de ajuste para los mismos, comprobando su precisión y su posible aplicación a fenómenos reales. Las contribuciones comentadas se han publicado o presentado en congresos internacionales según se explica detalladamente en el Capítulo Objetivos y en el desarrollo de los contenidos de esta memoria de doctorado.

Agradecimientos

El desarrollo de esta memoria así como la posibilidad de que distintos resultados que en ella se incluyen hayan podido presentarte en congresos, ha sido posible gracias al apoyo institucional y los siguientes recursos de financiación a lo largo del tiempo.

- Grupo de investigación *Modelización y Predicción con Datos Funcionales* (FQM-307) financiado por la Consejería de Educación, Ciencia y

Tecnología de la Junta de Andalucía.

- Grupo de investigación *Estadística Teórica y Aplicada e Investigación Operativa* FQM-245) financiado por la Consejería de Educación, Ciencia y Tecnología de la Junta de Andalucía.
- Proyecto de investigación *Desarrollo de metodologías funcionales para el análisis de datos estadísticos* (MTM2010-20502) financiado por el Plan Nacional I+D+I del Ministerio de Ciencia e Innovación.
- Proyecto de investigación *Avances en modelización con datos funcionales. Aplicación en Análisis de tiempos de vida* (MTM2013-47929-P) financiado por el Ministerio de Economía y Competitividad.

Me gustaría agradecer de un modo especial a mis directoras de tesis, la doctora Paula Rodríguez Bouzas y la doctora Nuria Ruiz Fuentes. Agradecer, porque durante todo este tiempo en el que he tenido altos y bajos, nunca me han abandonado, ayudándome en cada paso, consiguiendo que a pesar de la distancia las sintiera cada día a mi lado, sabiendo siempre darme ánimos. De ellas he aprendido mucho, su disposición, su incansable capacidad de trabajo, su continuo esfuerzo, su inquietud investigadora siempre pensando en el siguiente paso, en qué pasaría si cambiamos esto o lo otro. En estos años, gracias a su cercanía y humanidad, han pasado de ser profesoras a amigas y de amigas a familia, porque me han abierto a parte de su mente y su conocimiento, su corazón, haciéndome partícipe de su día a día, porque han sido muchos días y muchos acontecimientos los que han acaecido en este tiempo, unos buenos y otros no tanto, pero aún así siempre tienen una sonrisa y una palabra amable. Gracias de corazón.

A mi familia agradecerles que siempre han creído en mí, por su cariño y apoyo incondicional, por no dejarme decir que no puedo y por estar ahí cada

vez que los necesito sin necesidad ni de pedirlo. A mi compañero de vida, por su paciencia, por su capacidad de motivación, por no dejarme caer, y por tantos sacrificios que ha ido haciendo durante todo este tiempo para que esto pudiera ser, a él le debo todas esas horas en las que he podido dedicarme en exclusiva a mis estudios mientras él hacía todo lo demás. A mi hija, por la paciencia, por ser tan maravillosa, por preocuparse de si necesito algo, y por esos besos que en los momentos bajos dan la vida, y porque aun siendo tan pequeña ha sabido comprender mi ausencia en tantos ratos de juego. Y a mi talismán incansable, porque sea la hora que sea ella siempre está a mi lado, sin hacer más ruido que un quedo ronroneo , sin pedir nada, solo dando cariño y compañía cuando más lo necesitas, gracias.

Granada, 15 de octubre de 2019

Objetivos

Los objetivos de esta memoria de doctorado se han ido apuntando de forma implícita en la Introducción, ahora se expondrán explícitamente. El objetivo genérico es el estudio de procesos de recuento doblemente estocásticos partiendo del de Cox, tanto en su componente descriptiva (estudio de sus estadísticos) como en cuanto a su inferencia (estimación y contraste de hipótesis) desde la perspectiva clásica de Snyder y Miller (1991) [117].

Por supuesto, no en todos los procesos abordados se ha podido llegar al mismo nivel de profundización; ya sea porque en algún momento se ha atendido a las necesidades que surgían a partir del trabajo con datos reales, como por la complejidad encontrada.

En cada capítulo se estudia un proceso de recuento doblemente estocástico y en el desarrollo de su presentación se incluyen las aportaciones originales en el punto apropiado. A continuación, resaltamos y detallamos los objetivos específicos capítulo a capítulo de forma sistemática.

Capítulo 1. Proceso de Cox

Estudiar el proceso de Cox y

1. definir el proceso tipo fase para proponer su uso como proceso intensidad del de Cox (Bouzas et al., 2019 [33]) y

2. construir un contraste de bondad de ajuste para un proceso de Cox con proceso intensidad conocido (Ruiz-Fuentes et al., ERCIM 2016 [106]).

Capítulo 2. Proceso de Cox compuesto

Estudiar el proceso de Cox compuesto y

1. aportar expresiones cerradas para los estadísticos de tiempo (Montes-Gijón et al., ERCIM 2016 [87]; Bouzas et al., 2019 [33]),
2. desarrollar estimaciones para los estadísticos de tiempo incluyendo una proposición que resuelve el problema que surge al estimar la probabilidad de la ocurrencia de un nuevo suceso en un cierto intervalo de tiempo (Bouzas et al., 2019 [33]) y
3. construir contrastes de bondad de ajuste para un proceso de Cox compuesto con proceso intensidad conocido o desconocido (Ruiz-Fuentes et al., ERCIM 2016 [106]).

Capítulo 3. Proceso auto-excitado

Estudiar el proceso auto-excitado y

1. obtener expresiones explícitas de estadísticos de recuento,
2. construir un estimador puntual para la intensidad condicionada al recuento,
3. conseguir estimadores para estadísticos de recuento y del proceso intensidad.

Los resultados que prueban la consecución de estos objetivos se presentaron en (Ruiz-Fuentes et al., ERCIM 2014 [105]).

Capítulo 4. Simulaciones y Aplicaciones

Simular procesos de Cox, compuestos y autoexcitados con distintas intensidades y aplicar los resultados de la memoria tanto a las simulaciones como a datos reales. Concretando detalles, los objetivos se pueden enumerar de la siguiente manera,

1. simular procesos de Cox con intensidades diferentes y añadir un espacio de marcas para conseguir un proceso de Cox compuesto, de ahí, extraer procesos con marcas específicas (Bouzas et al., ERCIM 2014 [30]).
2. simular un proceso tipo fase y luego usarlo como proceso intensidad de uno de recuento, en particular, de un proceso de Cox compuesto o un auto-excitado (Ruiz-Fuentes et al., ERCIM 2013 [104]),
3. aplicar las estimaciones presentadas en la memoria a los procesos anteriores (estimación del proceso media, moda, probabilidad de una nueva ocurrencia) probando su precisión (Montes-Gijón et al., ERCIM 2016 [87]),
4. aplicar el test de bondad de ajuste para intensidad conocida o desconocida y, en el segundo caso (contribución de esta memoria), probar su buen comportamiento (Ruiz-Fuentes et al., ERCIM 2016, [106]),
5. aplicar todo lo anterior a ejemplos de datos reales: puntos de cambio de tendencia en cotizaciones bursátiles (Ruiz-Fuentes et al., CMS 2011 [103]) y datos extremos en temperaturas (Ruiz-Fuentes et al., CEB 2017 [107], Bouzas et al., 2019 [33]),
6. simular procesos auto-excitados con diferentes intensidades y distintas formas de dependencia (Ruiz-Fuentes et al., ERCIM 2014 [105]),

7. aplicar las estimaciones desarrolladas en la memoria a los procesos auto-excitados simulados (estimación de la intensidad condicionada al recuento, de la media y de la media de la intensidad) (Ruiz-Fuentes et al., ERCIM 2014 [105]).

Con la consecución de estos objetivos específicos, como se verá a lo largo de la memoria y se resumirá en las conclusiones, se cree alcanzado el propósito inicial de profundizar en el estudio teórico de los procesos de recuento doblemente estocásticos y en su inferencia, así como su aplicación a datos reales.

Capítulo 1

Proceso de Cox

Existen muchas razones para generalizar los procesos de Poisson, pero si hubiera que escoger una, quizá la más relevante es la de poder contemplar la posible dependencia de la ocurrencia de los sucesos de un proceso puntual, de otro proceso externo. Es por ello que, en este capítulo se desarrolla en estudio de un proceso de Poisson en el que se ha aleatorizado su intensidad a partir de un proceso tal que las trayectorias del proceso intensidad resultante no son conocidas dado solo el proceso puntual. Habitualmente, se formaliza esa influencia modelizándola a través de la intensidad del proceso puntual, sin embargo, como veremos más adelante esto puede hacerse también de otras formas. El proceso puntual resultante de dicha aleatorización, es un proceso de Poisson condicionado, dado el proceso intensidad, y se denomina Proceso de Poisson doblemente estocástico.

El *proceso de Poisson doblemente estocástico* o también llamado *proceso de Cox*, debido a que fue David R. Cox, 1955 [40] quien primero dedicó un estudio en profundidad a este tipo de procesos. Como ya se ha comentado, este proceso es una generalización del proceso de Poisson, con la singularidad de que su intensidad de ocurrencia depende de un proceso externo, siendo por

tanto dicha intensidad, un proceso estocástico también. El proceso externo del cual depende el proceso intensidad recibe el nombre de *proceso información*, conocido éste, el proceso puntual es un proceso de Poisson, por lo que el proceso de Cox también se conoce como proceso de Poisson condicionado. La naturaleza estocástica de su intensidad, hace que el proceso de Cox sea más realista y flexible en la modelización de fenómenos puntuales reales. Además, este nuevo concepto de doble aleatoriedad del proceso constituye un gran cambio en los procesos puntuales, particularmente en los de Poisson.

Volviendo al proceso de Poisson, base del desarrollo del proceso de Cox, encontramos que el estudio más amplio y detallado tanto del proceso de Poisson como sus generalizaciones entre las que se encuentra el proceso compuesto o filtrado, fue publicado por Snyder y Miller, 1975 [116] y en la segunda edición de Snyder y Miller, 1991 [117, pág. 344]. Ya en su último capítulo se presenta el proceso de Cox y alguna de sus generalizaciones de modo resumido. Si bien el proceso de Cox es fácil de definir a partir del proceso de Poisson, a menudo las expresiones explícitas de sus estadísticos son intratables como los propios Snyder y Miller, 1991 [117] reconocen. Aunque sin duda, el principal problema que presenta el estudio de los procesos de Cox, es la estimación del proceso intensidad. Son muchos los métodos que abordan este problema computacional utilizando desde cadenas de Markov, Análisis de Componentes Principales, regresiones poissonianas, hasta incluso nuevos métodos que surgen como el REX de Hutchison, 2013 [68]. Quizá sea esta dificultad el principal motivo por el que los trabajos de investigación dedicados a estos procesos suelen ser muy restrictivos en cuanto a sus objetivos y sus ejemplos de simulación, de hecho las últimas aportaciones en la literatura estadística acerca de ese último tipo de procesos, son escasas y con información muy dispersa. Es por todo ello, por lo que se considera el proceso de

Cox interesante tanto para su estudio como para el perfeccionamiento de su inferencia.

1.1. Definición

A la hora de desarrollar el estudio del proceso de Cox, basaremos éste en la definición formal que hicieron Snyder y Miller, 1991 [117].

Sea $\{x(t) : t \geq t_0\}$ un proceso continuo a la izquierda, que representa al proceso externo influyente en la intensidad de ocurrencia de los puntos del proceso de recuento $\{N(t) : t \geq t_0\}$, el proceso de Poisson doblemente estocástico o proceso de Cox, en adelante PC, se define como sigue:

Definición 1.1.1 *Un proceso de recuento $\{N(t) : t \geq t_0\}$ es un PC con proceso intensidad $\{\lambda(t, x(t)) : t \geq t_0\}$ si para casi toda trayectoria del proceso información $\{x(t) : t \geq t_0\}$, $N(\cdot)$ es un proceso de Poisson con función intensidad $\{\lambda(t, x(t)) : t \geq t_0\}$.*

En otras palabras:

El proceso de recuento $\{N(t) : t \geq t_0\}$ es un PC con proceso intensidad $\{\lambda(t, x(t)) : t \geq t_0\}$, si condicionado a $\{x(t) : t \geq t_0\}$ es un proceso de Poisson con función de intensidad $\{\lambda(t, x(t)) : t \geq t_0\}$.

Esta definición no es la única que se puede encontrar en la literatura, Serfozo, 1972 [112] define el PC basándose en la media del proceso. Concretamente, lo define como un proceso de Poisson cuya media es un proceso estocástico. La definición es la siguiente:

Definición 1.1.2 Sea $\{\Lambda(t) : t \geq t_0\}$ un proceso estocástico no decreciente, real valuado y continuo a la derecha, definido en el espacio de probabilidad (Ω, F, P) con $\Lambda(0) = 0$ c.s. Sea \mathcal{G} el menor σ -álgebra en Ω donde cada $\Lambda(t)$ sea medible. El proceso que toma valores enteros no negativos $\{N(t) : t \geq t_0\}$ definido en el mismo espacio de probabilidad, se denomina proceso de Poisson doblemente estocástico, PC, con media $\Lambda(t)$ si $N(t)$ es un proceso de Poisson con función media $\Lambda(t)$ condicionado a \mathcal{G} . Es decir, si $\Lambda(0) = 0$ c.s. y dados $t_0 < t_1 < \dots < t_r$ y n_1, \dots, n_r se tiene

- $P[N_{t_1} - N_0 \leq n_1, \dots, N_{t_r} - N_{t_{r-1}} \leq n_r / \mathcal{G}] = \prod_{i=1}^r P[N_{t_i} - N_{t_{i-1}} \leq n_i / \mathcal{G}]$ c.s. para todo $t_0 \leq t_1 \leq \dots \leq t_r$ y $n = 0, 1, 2, \dots$
- $P[N_t - N_s \leq n / \mathcal{G}] = \frac{1}{n!} (\Lambda_t - \Lambda_s)^n \exp[-(\Lambda_t - \Lambda_s)]$ c.s.

donde $N_{t_i} = N(t_i)$.

Ambas deficiones del PC, tienen puntos de vista distintos, mientras que en la definición de Snyder y Miller, 1991 [117] se define el PC a partir de su condicionamiento al proceso información y a través de la intensidad, en la que realiza Serfozo, 1972 [112] se define en base a la media del proceso. Ambas perspectivas pueden ser muy útiles tanto en el estudio del PC como en su caracterización, y por ello a lo largo de esta memoria haremos uso de ambas indistintamente.

A partir de las definiciones se observa que en el caso particular en el que la intensidad del proceso, $\lambda(t)$, es una función determinista del tiempo, el PC queda reducido a un proceso de Poisson no homogéneo, o proceso de Poisson general, sin olvidar que la diferencia con el proceso de Poisson homogéneo es que en este último su intensidad es constante.

Manteniendo la doble aleatoriedad, el PC más simple que se puede encontrar, es aquel en el que el proceso externo, el proceso información $\{x(t) : t \geq t_0\}$, es un conjunto finito de variables aleatorias independientes del tiempo. Un ejemplo sencillo de la aplicación de estos PC, es la representación de procesos de Poisson superpuestos donde el número de procesos superpuestos cambia aleatoriamente sin ocurrencias simultáneas.

De modo más formal, se suponen k procesos de Poisson independientes con intensidades $\{\lambda_1, \dots, \lambda_k\}$, cada uno de ellos interviniendo en un cierto instante o no, según el proceso estocástico $\{\xi(t) = (\xi_1(t), \dots, \xi_k(t)) : t \geq 0\}$, donde $\xi_i(t) = 1$ o 0 dependiendo si el i -ésimo proceso de Poisson interviene o no. Entonces el número de sucesos que ocurren en el intervalo $(0, t]$ es un PC con intensidad $\lambda(t) = \sum_{i=1}^k \lambda_i \xi_i(t)$.

Al igual que ocurre con el proceso de Poisson, también existen generalizaciones del PC, como es el caso del *proceso de Cox compuesto* que es la extensión natural del proceso de Poisson compuesto, donde la intensidad es un proceso estocástico influenciado por el proceso información. Este proceso es un proceso puntual marcado verificando que el proceso puntual es de Poisson y que las marcas de ocurrencia son variables aleatorias idénticamente distribuidas e independientes del proceso de Poisson. El proceso de Cox compuesto se suele utilizar a la hora de modelizar fenómenos reales como el registro de terremotos, en este caso las distintas magnitudes de los seísmos serían las marcas del proceso.

Otra generalización del PC, y en este caso una extensión natural del proceso de Cox compuesto, vendría dada al incrementar la dimensión del dominio del proceso puntual, surgiendo así el *proceso de Cox multidimensional* o *multichannel*. Este proceso es un PC definido sobre un dominio espacio temporal, donde el dominio espacial consiste en regiones discretas sobre las

que se observa un PC temporal, mientras las localizaciones espaciales dentro de las regiones son ignoradas.

El proceso puntual en el cual los puntos tienen ambas coordenadas, tanto la espacial como la temporal, puede verse como un proceso marcado en el tiempo, con marcas continuas que indican la posición espacial de un punto. Este caso particular de proceso se denomina *proceso de Cox espacio-temporal*.

Otra generalización más del PC es el *Proceso de Cox filtrado* que es aquel proceso que se puede modelizar como respuesta a los procesos marcados. Algunas de estas generalizaciones del PC se desarrollarán con más detalle a lo largo de esta memoria.

Como ya se ha apuntado con anterioridad y como se verá especialmente al tratar la inferencia del PC, el objetivo de esta memoria es obtener resultados aplicables a cualquier PC, tanto en este capítulo, como para los diversos procesos tratados en los capítulos posteriores. A parte dichos resultados, se pretende que sean alcanzados sin apenas realizar ninguna asunción acerca de su estructura estocástica ni de su proceso intensidad, ni de ningún otro proceso, sino a partir de la observación de trayectorias muestrales de dichos procesos. Por esta razón no se han descrito los modelos concretos de las generalizaciones del PC, tan solo han sido comentadas. En esta memoria, los modelos con ciertas intensidades o medias con modelos particulares solo nos interesarán para desarrollar los ejemplos o para realizar las simulaciones que ilustrarán los cálculos de los métodos o resultados obtenidos.

En este punto, se hará una excepción con la presentación de un modelo en el siguiente apartado. Esta particularidad se debe a que es una propuesta de esta memoria y que jugará un papel significativo como caso de simulación más adelante. El modelo propuesto es el PC cuya intensidad es un proceso tipo fase.

PC con intensidad proceso tipo fase

Para el desarrollo de este modelo particular de PC, hemos introducido brevemente las distribuciones tipo fase. Las distribuciones tipo fase, tienen su origen más lejano en Erlang, 1909 [48], para posteriormente ser introducidas por Neuts, 1975 [89] y más tarde descritas con detalle en Neuts, 1981 [90]. Este tipo de distribuciones han sido utilizadas en diversos campos del conocimiento como teoría de colas, supervivencia y fiabilidad (Asmussen, 2003 [8], Dayar, 2005 [44], Bladt, 2005 [15], Fackrell, 2008 [50], Reineck, 2012 [98], He, 2014 [66], Tank, 2015 [121], Ruiz-Castro, 2016a [101], Ruiz-Castro, 2016b [102]).

Las distribuciones tipo fase son distribuciones definidas en la semirecta real positiva y que tienen dos características muy importantes, su generalidad y su tratabilidad. Esta clase de distribuciones son una generalización de la distribución exponencial que cuentan con estructura matricial, por este motivo los estadísticos asociados a ellas se pueden obtener mediante algoritmos implementados computacionalmente. Como casos particulares de las distribuciones tipo fase se pueden nombrar la distribución Erlang, la distribución Coxian y la distribución hiperexponencial entre otras.

Sin embargo, aún hay más, ya que cualquier distribución discreta no-negativa con soporte finito es una tipo fase (Neuts, 1975 [89]). Incluso, esta clase de distribuciones es densa en las distribuciones de probabilidad no-negativas continuas. Por ello, dada una distribución no-negativa siempre hay una distribución tipo fase que se aproxima a ella tanto como se desee. Para un estudio en profundidad de las propiedades de las distribuciones tipo y fase, y más concretamente de sus propiedades de aproximación se pueden consultar en Neuts, 1975 [89], Shi, 1996 [113] y Asmussen, 2003 [8].

Si recordamos, la intensidad de un proceso de Cox, y en general de cual-

quier proceso de recuento, es no-negativa, por tanto vistas las propiedades de las distribuciones tipo fase, resultan ser candidatas perfectas para modelizarla. De hecho, es muy habitual encontrar trabajos en los que proponen como modelo de la intensidad de algún proceso de recuento, tanto distribuciones exponenciales como combinaciones lineales de ellas que son más que casos particulares de distribuciones tipo fase. Algunos ejemplos de esto se pueden encontrar en Hawkes, 1971 [65], Møller, 2004 [86], Ogata, 2006 [93], Bauwens, 2009 [12], Asha, 2010 [9], Russell [109], Lewis, 2012 [79], Chen, 2013 [36], Lefebvre, 2015 [77], Bieniek, 2017 [14], Sepehrifar, 2017 [111], Signh, 2018 [115]. Sin embargo, la intensidad de un PC no es una simple variable aleatoria, que sería el caso más simple, sino que en realidad es un proceso estocástico. Por esta razón, en esta memoria se propone tomar como modelo para el proceso intensidad una distribución tipo fase en la que se han aleatorizado los parámetros, surgiendo así un nuevo proceso estocástico al que denominaremos *proceso tipo fase*. Y así, de esta manera, conseguimos un proceso de recuento doblemente estocástico, en este caso un PC, muy general y flexible.

En los próximos capítulos también se usará este modelo de intensidad para los procesos de recuento tratados en ellos.

Con el fin de dejar establecida la notación que se utilizará a lo largo de esta memoria, se definen a continuación de modo formal las distribuciones tipo fase.

Lo más habitual es encontrar en la literatura la definición de distribución tipo fase ligada a las cadenas de Markov, del siguiente modo

Una distribución tipo fase es la distribución del tiempo de absorción de una cadena de Markov finita de dimensión $p + 1$, donde un estado es absorbente y los p restantes son transitorios, como puede verse en Neuts, 1975 [89].

Independientemente de este punto de vista, una distribución tipo fase viene dada por un vector fila $\boldsymbol{\pi}$ de dimensión p y una matriz \mathbf{T} de dimensión $p \times p$. A este par formado por $(\boldsymbol{\pi}, \mathbf{T})$ se le denomina *representación de la distribución tipo fase*.

El vector $\boldsymbol{\pi}$ se suele interpretar como un vector de probabilidad inicial, que recoge la probabilidad de pasar de un estado a otro, mientras que la matriz \mathbf{T} puede verse como la probabilidad de transición en un paso entre los estados transitorios de la cadena de Markov para el caso discreto y como el generador infinitesimal entre los estados transitorios en el caso continuo, o simplemente como una matriz de subintensidades.

En base a esto, la definición formal de la distribución tipo fase es la siguiente.

Definición 1.1.3 *Una variable aleatoria X se distribuye según una distribución tipo fase y se denota como $X \rightsquigarrow PH(\boldsymbol{\pi}, \mathbf{T})$ si su función de distribución viene dada por*

$$F(t) = P[X \leq t] = 1 - \boldsymbol{\pi} e^{\mathbf{T}t} \mathbf{e} \equiv 1 - \boldsymbol{\pi} \left(\sum_{n=0}^{\infty} \frac{(\mathbf{T}t)^n}{n!} \right) \mathbf{e}, \quad t \geq 0 \quad (1.1)$$

donde \mathbf{e} es el vector columna unitario; $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_p)_{1 \times p}$ es el vector de probabilidades iniciales, $\pi_i = P[X_0 = i]$ para $i = 1, \dots, p$, tal que $\sum_{i=1}^p \pi_i = 1$ y $P[X_0 = p+1] = 0$; $\mathbf{T} \in \mathcal{M}_p(\mathbb{R})$ es una matriz de subintensidades, verificando que los elementos de su diagonal son negativos: $\forall i = 1, \dots, p, T_{ii} < 0$, todos los elementos fuera de la diagonal son no negativos: $\forall i, j = 1, \dots, p, i \neq j, T_{ij} \geq 0$, la suma de los elementos de sus filas es no positiva: $\forall i, j = 1, \dots, p, \sum_{j=1}^p T_{ij} \leq 0$ y que \mathbf{T} es invertible.

Para una representación $(\boldsymbol{\pi}, \mathbf{T})$ dada la definición 1.1.3, la función definida en la ecuación (1.1) es una función de distribución y su función de

densidad está dada por

$$f(t) = \frac{dF(t)}{dt} = \boldsymbol{\pi} e^{\mathbf{T}t} \mathbf{T}^0, \quad t \geq 0, \quad (1.2)$$

donde $\mathbf{T}^0 = -\mathbf{T}\mathbf{e}$.

No es objetivo de esta memoria hacer un estudio exhaustivo de las distribuciones tipo fase, por lo que en este punto recordaremos solamente los momentos de estas distribuciones para obtener la media y la varianza, ya que estas son las medidas que se necesitarán en los ejemplos de simulación.

El momento de orden n de la variable aleatoria $X \rightsquigarrow PH(\boldsymbol{\pi}, \mathbf{T})$ es

$$\mu_n = E[X^n] = n! \boldsymbol{\pi} (-\mathbf{T})^{-n} \mathbf{e} \quad (1.3)$$

Una de las ventajas que tienen las distribuciones tipo fase en su uso en modelos estocásticos, es que el momento de orden n tiene expresión matricial. A partir de la ecuación (1.3) se obtienen de modo inmediato las expresiones de la media y la varianza, siendo estas las siguientes.

$$E[X] = -\boldsymbol{\pi} \mathbf{T}^{-1} \mathbf{e} \quad (1.4)$$

y

$$Var(X) = 2\boldsymbol{\pi} \mathbf{T}^{-2} \mathbf{e} - (\boldsymbol{\pi} \mathbf{T}^{-1} \mathbf{e})^2 \quad (1.5)$$

Hasta aquí se ha recogido la información básica de una distribución tipo fase, pero ha de recordarse que en esta memoria se propone el proceso tipo fase como modelo de intensidad, por lo que a continuación se define formalmente.

Definición 1.1.4 *Un proceso tipo fase es un proceso estocástico con representación $(\boldsymbol{\pi}, \mathbf{T})$, donde*

1. $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_p)_{1 \times p}$ y $\mathbf{T} = (\mathbf{T}_{ij})_{p \times p}$
2. siendo los parámetros $\boldsymbol{\pi}$ y/o \mathbf{T} variables aleatorias.
3. para cada realización de la variable aleatoria, su distribución es una distribución tipo fase.

El último término de la definición impone sobre las variables $\boldsymbol{\pi}$ y \mathbf{T} las propiedades y restricciones de las variables tipo fase que se describieron anteriormente en esta sección. Definido el proceso tipo fase, es fácil observar que si se utiliza como modelo de intensidad se obtiene un proceso intensidad muy general, sin necesidad de restricciones y a la vez muy flexible, por lo que se tiene un proceso doblemente estocástico muy general a su vez. Este nuevo proceso no solo puede utilizarse con el PC, sino que puede modelizar la intensidad de cualquier proceso de recuento de cualquier campo de estudio, por lo que se convierte en una herramienta de gran interés y utilidad.

Para una mejor ilustración, veamos un ejemplo de un PC con distribución tipo fase. La intensidad del proceso tiene la representación que aparece en la ecuación (1.6) y en la figura (1.1) se puede ver una trayectoria muestral.

$$(\boldsymbol{\pi}, \mathbf{T}(\beta_1, \beta_2, \beta_3)) \quad \text{donde} \quad (1.6)$$

$$\boldsymbol{\pi} = (0.2, 0.4, 0.4) \quad \text{y} \quad \mathbf{T}(\beta_1, \beta_2, \beta_3) = \begin{pmatrix} -1/\beta_1 & 0 & 0 \\ 0 & -1/\beta_2 & 0 \\ 0 & 0 & -1/\beta_3 \end{pmatrix}$$

con $\beta_1 \rightsquigarrow \mathcal{B}(10, 0.8) + 1$, $\beta_2 \rightsquigarrow \mathcal{B}(20, 0.5) + 1$ y $\beta_3 \rightsquigarrow \mathcal{B}(1, 0.5) + 1$.

Este ejemplo puede ampliarse, e incluso reinterpretarse, como se verá en el capítulo 4 de esta memoria dedicado a simulaciones.

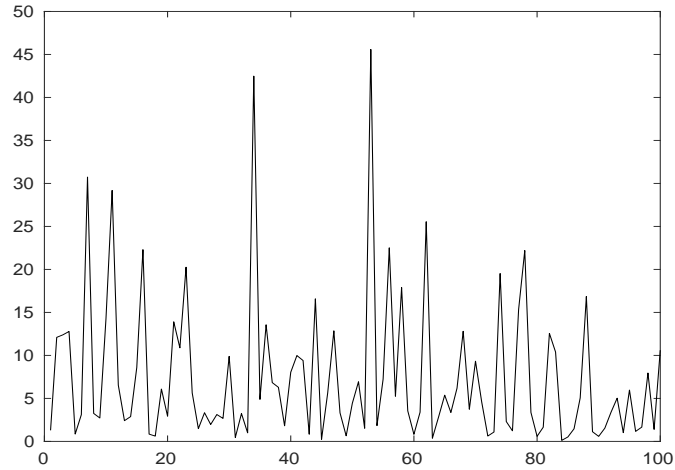


Figura 1.1: Trayectoria muestral de la intensidad como un proceso tipo fase.

1.2. Estadísticos de recuento

Una vez que ya se ha definido el PC, en esta sección se pasa a formalizar las expresiones de diversas características asociadas a dicho proceso y de las cuales se obtienen varias medidas y momentos.

Sea $\{N(t) : t \geq t_0\}$ un PC con proceso intensidad $\{\lambda(t, x(t)) : t \geq t_0\}$, donde $\{x(t) : t \geq t_0\}$ es el proceso información. Los estadísticos de recuento interés de $N(t)$ incluyen varios momentos, así como la función masa de probabilidad.

Entre los distintos métodos existentes a la hora de determinar dichos estadísticos, cabe destacar el método de condicionamiento desarrollado en Snyder y Miller, 1991 [117], y que está basado en que dado el proceso de información, $\{N(t) : t \geq t_0\}$ es un proceso de Poisson condicionado.

1.2.1. Función masa de probabilidad

En primer lugar veamos la conocida función masa de probabilidad, que no es más que la probabilidad de que el número de puntos del proceso que ocurren en el intervalo $[t_0, t)$ sea n y se define como

$$P[N(t) = n] = E \{P[N(t) = n/x(\sigma) : t_0 \leq \sigma < t]\} = E \left\{ \frac{1}{n!} \left(\int_{t_0}^t \lambda(\sigma, x(\sigma)) d\sigma \right)^n \exp \left(- \int_{t_0}^t \lambda(\sigma, x(\sigma)) d\sigma \right) \right\} \quad (1.7)$$

para $n = 0, 1, 2, \dots$

Aunque a partir del condicionamiento al proceso información se llega de un modo explícito a la expresión de la función masa de probabilidad, el hecho de encontrar una expresión para un caso concreto no es inmediato, sino que lo habitual es que sea muy difícil o incluso imposible de determinar completamente. Uno de los pocos ejemplos en los que se puede llegar a una expresión para su cálculo se encuentra en Snyder y Miller, 1991 [117, pág. 344] donde toman un proceso de Poisson a escala aleatoria, y obtienen una expresión de la función masa de probabilidad, también pueden encontrarse más ejemplos en Jiménez, 1991 [70], Bouzas, 1999 [18] y Bouzas et al. 2006a [21] en el que se desarrollan los estadísticos de recuento de un PC con intensidad un proceso de banda estrecha.

A la hora de abordar la función masa de probabilidad, podemos encontrarnos con que no se tiene conocimiento de la intensidad del PC, pero si de su media dado el proceso información, por lo que sería interesante su desarrollo desde este punto de vista.

Sea $\Lambda(t, x(t))$ la función paramétrica del proceso de Poisson condicionado, siempre que $\Lambda(t, x(t))$ sea absolutamente continua, lo que supone una débil restricción; en ese caso $\Lambda(t, x(t)) = \int_{t_0}^t \lambda(\sigma, x(\sigma)) d\sigma$ es la media del PC.

Dada esta relación, la función paramétrica es también un proceso influenciado por el proceso información, luego, puede reescribirse la función masa de probabilidad como

$$P[N(t) = n] = \frac{1}{n!} E \left\{ \Lambda(t, x(t))^n e^{-\Lambda(t, x(t))} \right\} = \frac{1}{n!} G_{\Lambda(t, x(t))}^n(-1) \quad (1.8)$$

donde $G_{\Lambda(t, x(t))}$ es la función generatriz de momentos de $\Lambda(t, x(t))$.

Esta nueva expresión de la función masa de probabilidad basada en la media $\Lambda(t, x(t))$ del proceso $N(t)$, pone de manifiesto el interés que esta memoria dedica a la definición de Serfozo, 1972 [112], en la que se basa su caracterización.

1.2.2. Función característica

Al igual que cualquier proceso estocástico el PC se caracteriza a partir de su función característica. Para definirla, se utiliza de nuevo el método de condicionamiento de Snyder y Miller, 1991 [117], obteniéndose que la función característica de un PC $N(t)$ con proceso de información $\{x(t); t \geq t_0\}$, se define como

$$M_{N(t)}(iu) = E \left\{ \exp \left[(e^{iu} - 1) \int_{t_0}^t \lambda(\sigma, x(\sigma)) d\sigma \right] \right\} \quad (1.9)$$

A parte de la expresión dada, la función característica de un PC también puede escribirse en función de la media del proceso o función paramétrica $\Lambda(t, x(t))$, del siguiente modo:

$$M_{N(t)}(iu) = E \left\{ \exp \left[(e^{iu} - 1) \Lambda(t, x(t)) \right] \right\} = M_{\lambda}(e^{iu} - 1) \quad (1.10)$$

Aunque se tengan estas dos expresiones de la función característica, de forma similar a lo que ocurría con la función masa de probabilidad, su evaluación

no siempre es posible. Resulta fácil observar que a partir de ella resulta que

$$E[N(t)] = E[\Lambda(t, x(t))] \quad \text{y} \quad \text{Var}[N(t)] = \text{Var}[\lambda(t, x(t))] + E[\lambda(t, x(t))] \quad (1.11)$$

y por tanto queda demostrado algo asumido hasta ahora debido a la asociación del PC con el proceso de Poisson. Esto es, que condicionado al proceso información, Λ es la media del PC.

Para el caso concreto en el que la intensidad del PC $\lambda(t, x(t))$ sea gaussiana, se puede encontrar en Valderrama et al., 1995 [122] una expresión evaluable para la función característica enunciada en el siguiente teorema.

Teorema 1.2.1 *La función característica de un proceso de Cox $\{N(t) : t \geq t_0\}$ con intensidad aleatoria gaussiana $\lambda(t, x(t))$ viene dada por*

$$M_{N(t)}(iu) = \exp \left\{ (e^{iu} - 1)^2 \int_{t_0}^t \int_{t_0}^s R_\lambda(v, s) dv ds + (e^{iu} - 1) \int_{t_0}^t m_\lambda(s) ds \right\} \quad (1.12)$$

donde $m_\lambda(t)$ y $R_\lambda(t)$ son la media y función de covarianza respectivamente, de la intensidad.

La demostración de este teorema expuesta en el citado artículo, tiene por base fundamental el desarrollo en serie de Karhunen-Loève que se hace de la intensidad $\lambda(t, x(t))$.

Otro ejemplo interesante para el cual es posible calcular expresiones explícitas tanto para la función característica como para algunos estadísticos, es el de un PC cuya media, $\Lambda(t, x(t))$ se distribuye mediante una distribución Normal truncada. El desarrollo de este caso se puede encontrar en Bouzas et al., 1999 [18].

1.2.3. Funcional característico

El funcional característico de una variable aleatoria fue introducido por primera vez por Andrej Kolmogorov, 1935 [73], como una generalización de la función característica. El funcional característico se define para medidas aleatorias generales en Daley y Vere-Jones, 1988 [41], o para procesos de Poisson en Snyder y Miller, 1991 [117], pero no para procesos de Cox.

Dada una variable aleatoria X en un espacio lineal L , se define el funcional característico de X como:

$$\Phi_X(iv) = E[e^{i\langle X, v \rangle}], \quad (1.13)$$

siendo v una función real valuada.

Basada en la definición para una medida aleatoria η encontrada en Daley y Vere Jones, 1988 [41],

$$\Phi_\eta(iv) = E \left[\exp \left\{ i \int v(\sigma) \eta(d(\sigma)) \right\} \right] \quad (1.14)$$

La definición del funcional característico para un PC se encuentra en Bouzas et al., 2006a [21].

Definición 1.2.2 *Funcional característico de un proceso de Cox*

Se define el funcional característico de un PC, como

$$\Phi_N(iv) \equiv \exp \left[i \int_0^T v(\sigma) N(d\sigma) \right] \quad (1.15)$$

siendo v una función real valuada, y la integral es una integral de recuento con evaluación

$$\int_0^T v(\sigma) N(d\sigma) = \sum_{i=1}^{N(T)} v(w_i)$$

donde w_i son los tiempos de ocurrencia del PC. Utilizando el método de condicionamiento se obtiene

$$\begin{aligned}\Phi_N(iv) &= E \left\{ \exp \left[i \int_0^T v(\sigma) N(d\sigma) / x(\sigma) : t_0 \leq \sigma < t \right] \right\} \\ &= E \left\{ \exp \left[\int_0^T \lambda(\sigma, x(\sigma)) (e^{iv(\sigma)} - 1) d\sigma \right] \right\}\end{aligned}\quad (1.16)$$

Para un estudio en mayor profundidad de este tipo de integrales, se puede ver el cálculo estocástico de Itô en Snyder y Miller, 1991 [117] o en Kloeden y Platen, 1999 [72].

Como se ha mencionado anteriormente, en la mayoría de los casos los estadísticos de un PC son intratables, es por ello interesante encontrar un modo para su cálculo. El funcional característico de un PC, proporciona la distribución finito-dimensional del proceso, así como sus momentos y permite determinar la distribución conjunta de un conjunto finito de variables del proceso. Diferentes funciones $v(\sigma)$ en la definición de funcional característico, permiten el cálculo de las funciones características de variables relacionadas con el PC como pueden ser la de los incrementos o las de los vectores finito-dimensionales que se muestran en los siguientes casos.

Caso 1.2.3 *Función característica de los incrementos*

Considerando como función real valuada $v(\sigma)$ la siguiente función

$$v(\sigma) = \begin{cases} 0, & 0 \leq \sigma < t_1; \\ \alpha, & t_1 \leq \sigma < t_2; \\ 0, & t_2 \leq \sigma < T. \end{cases} \quad 0 < t_1 < t_2 < T$$

se tiene que la expresión para el funcional característico (1.15) resulta

$$\Phi_N(iv) = E\{\exp[i\alpha N(t_1, t_2)]\} \equiv M_{N(t_1, t_2)}$$

que es la función característica del incremento del proceso de Cox entre t_1 y t_2 . Obteniéndose también que

$$\Phi_N(iv) = E \left\{ \exp \left[(e^{i\alpha} - 1) \int_{t_1}^{t_2} \lambda(\sigma, x(\sigma)) d\sigma \right] \right\}$$

o bien en términos de la función paramétrica o proceso media

$$\Phi_N(iv) = E \left\{ \exp \left[(e^{i\alpha} - 1) (\Lambda(t_2) - \Lambda(t_1)) \right] \right\} = M_{\Lambda(t_2) - \Lambda(t_1)}(e^{i\alpha} - 1),$$

siendo $M_{\Lambda(t_2) - \Lambda(t_1)}$ la función característica de $\Lambda(t_2) - \Lambda(t_1)$.

Caso 1.2.4 Función característica de vectores aleatorios finito-dimensionales

Se considera ahora la siguiente función $v(\sigma)$

$$v(\sigma) = \begin{cases} \alpha_1 + \alpha_2 + \dots + \alpha_m, & 0 \leq \sigma < t_1; \\ \alpha_2 + \dots + \alpha_m, & t_1 \leq \sigma < t_2; \\ \vdots & \vdots \\ \alpha_m, & t_{m-1} \leq \sigma < t_m; \\ 0, & t_m \leq \sigma < T. \end{cases}$$

donde $0 < t_1 < t_2 < \dots < t_m < T$.

A partir de la definición de funcional característico (1.15), se obtiene:

$$\begin{aligned} \Phi_N(iv) &= \\ E\{ \exp[&i(\alpha_1 + \dots + \alpha_m)N(t_1) + i(\alpha_2 + \dots + \alpha_m)(N(t_2) - N(t_1)) + \\ &\dots + i\alpha_m(N(t_m) - N(t_{m-1}))] \} = \\ E\{ \exp[&i\alpha_1 N(t_1) + \dots + i\alpha_m N(t_m)] \} \equiv M_{N(t_1), \dots, N(t_m)}(i\alpha_1, \dots, i\alpha_m) \end{aligned}$$

donde $M_{N(t_1), \dots, N(t_m)}(i\alpha_1, \dots, i\alpha_m)$ es la función característica conjunta de $(N(t_1), \dots, N(t_m))$. De esta expresión también se obtiene

$$\begin{aligned} \Phi_N(iv) &= \\ E \left\{ \exp \left[(e^{i(\alpha_1 + \dots + \alpha_m)} - 1) \int_0^{t_1} \lambda(\sigma) d\sigma + \dots + (e^{i\alpha_m} - 1) \int_{t_{m-1}}^{t_m} \lambda(\sigma) d\sigma \right] \right\} &= \\ E \left\{ \left[(e^{i(\alpha_1 + \dots + \alpha_m)} - e^{i(\alpha_2 + \dots + \alpha_m)}) \Lambda(t_1) + \dots + (e^{i\alpha_m} - 1) \Lambda(t_m) \right] \right\} &= \\ M_{\Lambda(t_1), \dots, \lambda(t_m)}(e^{i(\alpha_1 + \dots + \alpha_m)} - e^{i(\alpha_2 + \dots + \alpha_m)}, \dots, e^{i\alpha_m} - 1). \end{aligned}$$

Por simplificar la notación, se ha omitido la dependencia de la media del proceso información.

Para el caso particular en el que el número de tiempos $m = 2$, se obtiene la función característica de las distribuciones bidimensionales del proceso de Cox a partir de la cual se pueden calcular momentos de $(N(t_1), N(t_2))$ y su función de covarianza:

$$R_N(t_1, t_2) = R_\Lambda(t_1, t_2) + E[\Lambda(t_1)]$$

donde $R_\Lambda(t_1, t_2)$ es la covarianza del proceso media. Ver Bouzas et al., 2006a [21] para consultar las demostraciones.

1.2.4. Moda

Por definición, la moda en un instante fijo de tiempo, es el valor que el proceso toma con mayor probabilidad en dicho instante. El cálculo de este estadístico no es viable a partir de la función característica. No siendo usual en la literatura, el interés surgió en el entorno del PC compuesto, por lo tanto, más adelante en esta memoria, concretamente en el capítulo al PC compuesto se demostrará cuál es la expresión de la moda en términos más generales.

1.3. Estadísticos de tiempo

A parte de para el estudio y desarrollo de los estadísticos de recuento como el cálculo de la función masa de probabilidad y la función característica de un PC, el método de condicionamiento también se utiliza para la obtención de una expresión de los estadísticos de tiempo, estadísticos que estudian la localización y los espacios entre los puntos, los instantes de ocurrencia.

Se llama $p_W^{(n)}(w)$ a la función de densidad conjunta de los n primeros tiempos de ocurrencia w_1, \dots, w_n . Por la definición de densidad conjunta de un proceso de Poisson y haciendo uso del método de condicionamiento, se define la *función de densidad conjunta* de un PC como

$$p_W^{(n)}(w) = E \left\{ \prod_{i=1}^n \lambda(w_i, x(w_i)) \exp \left[- \int_{t_0}^{w_n} \lambda(\sigma, x(\sigma)) d\sigma \right] \right\}$$

con $t_0 \leq w_1 \leq w_2 \leq \dots \leq w_n$.

Haciendo uso nuevamente del método de condicionamiento y la definición de la función de densidad muestral de un proceso de Poisson, se define ahora la función de densidad muestral de un PC como

$$P[\{N(t); t_0 \leq \sigma < T\}] = E \left\{ \exp \left[- \int_{t_0}^T \lambda(\sigma, x(\sigma)) d\sigma + \int_{t_0}^T \ln \lambda(\sigma, x(\sigma)) N(d\sigma) \right] \right\} \quad (1.17)$$

donde la segunda integral es una integral de recuento.

El PC será generalizado por el proceso de Cox compuesto en el capítulo 2 y será entonces cuando se presentará un estudio de los estadísticos de tiempo más amplio ya que es una de las aportaciones de esta memoria.

1.4. Inferencia

Existe mucha literatura en la que se puede encontrar una gran cantidad de información acerca de procesos puntuales, sobre su estimación y sobre la predicción de su proceso intensidad, casi siempre asumiendo un conocimiento estadístico previo sobre dicha intensidad, como pueden ser su distribución o momentos. Sin embargo, no se encuentra mucho trabajo hecho sobre estimación del PC, a pesar de ser un proceso de gran interés dada su versatilidad. Se pueden encontrar estudios acerca del PC y sus generalizaciones desde diversos puntos de vista como puede ser Chiang et al., 2005 [38], Ho y Stoyan, 2008 [67], Fernández-Alcalá et al., 2009 [51], Fernández-Alcalá, 2010 [52], Myllymaki y Pettinen, 2010 [84], Zhang y Kou, 2010 [125], Zhao et al., 2011 [127], entre otros, aunque la mayoría de ellos se dedican a la estimación y pocos a la predicción.

A lo largo de esta sección se procederá al estudio de la inferencia sobre el PC, y para ello se hará uso del Análisis en Componentes Principales, del Análisis en Componentes Principales Funcional y de modelos de Predicción en Componentes Principales

El *Análisis en Componentes Principales (ACP)* es una técnica de análisis multivariante que tiene como finalidad reducir la dimensión del problema mediante la construcción de transformaciones lineales de las variables, siendo dichas transformaciones incorreladas y de varianza máxima. Además el *Análisis en Componentes Principales Funcional (ACPF)* es una técnica que permite estimar un proceso estocástico cualquiera para después predecir mediante los *modelos de predicción en componentes principales (PCP)*, desarrollado por Aguilera et al., 1997 [4], a partir de trayectorias muestrales del proceso y asumiendo pocas restricciones teóricas. Es por este motivo por el

que resultan una herramienta de gran utilidad a la hora de estimar y predecir el PC, pretendiendo hacerlo siempre a partir de las observaciones del proceso y sin necesidad de imponer una estructura prefijada al proceso intensidad o al proceso media.

Tanto la intensidad como la media del PC y del PC compuesto, son procesos estocásticos, por lo que se pueden estimar y predecir, para ello utilizaremos modelos PCP como se acaba de explicar. Haciendo uso de esta técnica se tiene la posibilidad de estimar a parte de la intensidad o la media, importantes en sí mismas para muchos fenómenos reales, el propio PC y sus estadísticos, dado que todos ellos se expresan en función de alguna de ellas o incluso de ambas.

1.4.1. Inferencia del proceso media

- *Estimación*

Lo primero que hay que abordar a la hora de hacer inferencia sobre un PC, es dar una estimación estocástica de su proceso media, y para ello $\{N(t); t \geq t_0\}$ debe estimarse puntualmente la media para después poder reconstruir sus trayectorias funcionales y finalmente terminar aplicando ACPF. En el artículo Bouzas et al., 2006c [23] se propone una aproximación al ACPF mediante la interpolación cúbica monótona de las trayectorias para así conservar la propiedad teórica de no decrecimiento. Puesto que ésta reconstrucción no es una de las que se encuentran habitualmente en la literatura del ACPF, se recuerda el procedimiento a continuación (consultar el artículo para desarrollar los detalles).

En primer lugar se estiman k funciones muestrales $\Lambda_w(t)$ en un con-

junto finito de instantes de tiempo a partir de k funciones muestrales independientes del PC $\{N_{wi}(t) : w = 1, \dots, k\}$ observadas en el intervalo de tiempo $[t_0, t_0 + rT]$. Para realizar dicha estimación, cada una de las funciones muestrales observadas se subdivide en r subtrayectorias de amplitud T que, dado que el PC tiene incrementos independientes, pueden reinterpretarse como observadas en el mismo subintervalo $[T_0, T_1) \equiv [t_0, t_0 + T)$. Se notarán las subtrayectorias como

$$N_{wi}(t) = N_w(t + (i - 1)T) - N_w(t_0 + (i - 1)T), \quad t \in [t_0, t_0 + T)$$

con $w = 1, \dots, k$ e $i = 1, \dots, r$.

Tomando la nueva partición

$$t_0 = t_0 < t_1 < \dots < t_j < \dots < t_p = t_0 + T \quad (1.18)$$

el estimador puntual del proceso media en esos instantes de tiempo es

$$\hat{\Lambda}_w(t_j) = \frac{\sum_{i=1}^r N_{wi}(t_j)}{r}, \quad (1.19)$$

para $j = 1, \dots, p$ y $w = 1, \dots, k$. Por tanto se obtienen valores de estimaciones puntuales de $\Lambda(t, x(t))$ en $p + 1$ instantes de tiempo en el intervalo $[t_0, t_0 + T)$ para k trayectorias.

Después de estimar las k funciones muestrales, el siguiente paso es la reconstrucción de la forma funcional de las trayectorias estimadas de forma discreta. Los métodos habituales que se pueden encontrar en la literatura del ACP, no permiten asegurar que esa reconstrucción conserve la monotonía que deben tener las trayectorias de $\Lambda(t, x(t))$. Por esta razón en Bouzas et al., 2006c [23] se adapta el ACP a las necesidades particulares del PC. Basándose en la interpolación cúbica monótona a trozos de Fritsch y Carlson, 1980 [53], el polinomio de

interpolación de los datos $\{(t_0, \Lambda_w(t_0)), \dots, (t_p, \Lambda_w(t_p))\}$ es

$$p_{wj} = \Lambda_w(t_j)H_1(t) + \Lambda_w(t_{j+1})H_2(t) + d_{wj}H_3(t) + d_{wj+1}H_4(t), \quad \begin{array}{l} t \in [t_j, t_{j+1}] \\ j = 0, \dots, p-1 \end{array} \quad (1.20)$$

donde $d_{wj} = \left. \frac{dp_{wj}(t)}{dt} \right|_{t=t_j}$, $\left. \frac{dp_{wj}(t+1)}{dt} \right|_{t=t_{j+1}}$ y las funciones $H_s(t)$ forman la base usual de Hermite en el intervalo $[t_j, t_{j+1}]$. Para poder aplicar el ACPF, es necesario que las interpolaciones estén expresadas en todo el intervalo y no a trozos. Luego, reajustando la base del espacio de funciones se puede escribir que la interpolación monótona de cada trayectoria es

$$I\Lambda_w(t) = \sum_{j=0}^p \Lambda_w(t_j)\Phi_j(t) + \sum_{j=0}^p d_{wj}\Psi_j(t), \quad t \in [t_0, t_0 + T], \quad w = 1, \dots, k. \quad (1.21)$$

donde las funciones $\Phi_s(t)$ y $\Psi_s(t)$ forman la base usual de Lagrange de splines cúbicos. Para unificar notación, se puede escribir

$$I\Lambda(t) = AB(t), \quad t \in [t_0, t_0 + T],$$

con

$$\begin{aligned} I\Lambda(t) &= (I\Lambda_1(t), \dots, I\Lambda_k(t))^T, \\ B(t) &= (\Phi_0(t), \dots, \Phi_p(t), \Psi_0(t), \dots, \Psi_p(t))^T, \end{aligned} \quad (1.22)$$

$$A = \begin{pmatrix} \Lambda_1(t_0) & \dots & \Lambda_1(t_p) & d_{10} & \dots & d_{1p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Lambda_k(t_0) & \dots & \Lambda_k(t_p) & d_{k0} & \dots & d_{kp} \end{pmatrix}, \quad (1.23)$$

donde T denota trasposición. Para lograr una mayor unificación se renombran los elementos de $B(t)$ y de A como

$$B(t) = (B_1(t), \dots, B_{p+1}(t), B_{p+2}, \dots, B_{2(p+1)})^T, \quad (1.24)$$

$$A = (a_{wl}) \quad w = 1, \dots, K \quad l = 1, \dots, 2(p+1). \quad (1.25)$$

Por lo que el polinomio (1.21) puede escribirse

$$I\Lambda_w(t) = \sum_{l=1}^{2(p+1)} a_{wl} B_l(t); \quad t \in [t_0, t_0 + T), \quad w = 1, \dots, k. \quad (1.26)$$

Una vez reconstruidas las trayectorias funcionales del proceso media, se aplica la técnica de ACPF. La reducción de la dimensión es muy importante puesto que el número de observaciones muestrales, $p+1$, se ha duplicado, $2(p+1)$, tras el uso de la interpolación cúbica monótona. Descritos de un modo breve, los pasos a seguir para adaptar la metodología del ACPF serían los siguientes.

Considérese el proceso media interpolado centrado

$$\overline{I\Lambda}(t) = I\Lambda(t) - \mu_{I\Lambda}(t) = (A - \overline{A})B(t)$$

donde $\overline{A} = (\overline{a}_{wl})$ con elementos $\overline{a}_{wl} = \frac{1}{k} \sum_{w=1}^k a_{wl}$ ($l = 1, \dots, 2(p+1)$); $w = 1, \dots, k$) y $\mu_{I\Lambda}(t) = \overline{A}B(t)$.

Sea $(A - \overline{A})$ la matriz cuyas filas son los coeficientes de cada trayectoria del proceso centrado con respecto a la base $\langle B_1(t), \dots, B_{2(p+1)} \rangle$ y \mathbb{P} la matriz cuyas componentes son los productos interiores usuales entre las funciones de la base $\langle B_i, B_j \rangle_u = \int_{t_0}^{t_p} B_i(t) B_j(t) dt$. Entonces, el ACPF de la trayectoria interpolada $\overline{I\Lambda}_w(t)$ en el espacio generado por la base $\langle B_1(t), \dots, B_{2(p+1)} \rangle$ con respecto a la métrica usual en

$L^2[t_0, t_p]$ es equivalente al ACP multivariante de la matriz $(A - \bar{A})\mathbb{P}^{1/2}$ con respecto al producto interno usual en $\mathbb{R}^{2(p+1)}$. Una vez calculados los autovalores g_j de la matriz de covarianza $(A - \bar{A})\mathbb{P}^{1/2}$, las funciones muestrales del proceso interpolado $\bar{I}\bar{\Lambda}(t)$ pueden representarse en términos de sus componentes principales como

$$\bar{I}\bar{\Lambda}_w(t) = \sum_{j=1}^{2(p+1)} \zeta_{wj} f_j(t), \quad w = 1, \dots, k,$$

donde $f_j(t)$ son las autofunciones de la covarianza muestral del proceso media dadas por

$$f_j(t) = \sum_{l=1}^{2(p+1)} f_{lj} B_l(t)$$

con el vector de coeficientes $f_j = \mathbb{P}^{-1/2} g_j$ y

$$\zeta_{wj} = \int_{t_0}^{t_p} \bar{I}\bar{\Lambda}_w(t) f_j(t) dt = (A_w - \bar{A}_w) \mathbb{P}^{1/2} g_j,$$

donde $(A_w - \bar{A}_w)$ es la fila w -ésima de la matriz $(A - \bar{A})$. Por tanto, el proceso media se puede aproximar mediante una descomposición ortogonal truncada de sus componentes principales notada por $\Lambda^q(t)$

$$\Lambda^q(t) \simeq I\Lambda^q(t) = \mu_{i\Lambda}(t) + \sum_{j=1}^p \zeta_j f_j(t). \quad (1.27)$$

Por tanto, la dimensión $2(p+1)$ de las observaciones muestrales se ha reducido a q con una proporción de varianza explicada del proceso media tan próxima a 1 como sea posible dada por $\frac{\sum_{j=1}^q \lambda_j}{\sum_{j=1}^{2(p+1)} \lambda_j}$ donde λ_j es la varianza de la j -ésima componente principal ζ_j dada por el j -ésimo autovalor de la matriz de covarianzas $(A - \bar{A})\mathbb{P}^{1/2}$ asociado al j -ésimo autovector g_j . En Bouzas et al., 2006c [23] se comprueba el pequeño error de estimación que se consigue con el método propuesto mediante la simulación y posterior estudio de distintos ejemplos de PC

con intensidades distintas. Esta estimación de la media de un PC se puede extrapolar a la estimación del proceso media de un PC compuesto con marcas específicas.

■ *Predicción*

Como ya se comentó anteriormente, no es habitual encontrar literatura referente a la predicción del PC, por ello consideramos que interesante su estudio. Para ello, tras la estimación del proceso media en el apartado anterior, se considera la predicción del proceso media de un PC en un intervalo futuro $(T_1; T_2]$ a partir de una trayectoria observada de forma discreta en el intervalo del pasado $[T_0; T_1]$. Para realizar dicha predicción, se estudia el modelo a partir de trayectorias observadas en el intervalo $[T_0; T_2]$ y su comportamiento en el intervalo $(T_1; T_2]$ a partir del conocimiento que se tiene en el intervalo $[T_0; T_1]$. Si se aplica ACPF al proceso media en sendos intervalos, de igual modo al explicado en el apartado de estimación, las descomposiciones truncadas en función de sus componentes principales que se obtienen son:

$$\begin{aligned}\widehat{\Lambda^1}(t) &= \Lambda^{q_1}(t) = \mu_{\Lambda}^1(t) + \sum_{j=1}^{q_1} \zeta_j f_j(t); & t \in [t_0 = T_0, T_1] \\ \widehat{\Lambda^2}(s) &= \Lambda^{q_2}(s) = \mu_{\Lambda}^2(s) + \sum_{j=1}^{q_2} \eta_j g_j(s); & s \in (T_1, T_2)\end{aligned}\quad (1.28)$$

Denotemos por $\tilde{\eta}_j^{p_j} = \sum_{i=1}^{p_j} b_i^j \xi_i$ al estimador de η_j , $j = 1, \dots, q_2$ en función de las p_j componentes principales de ξ_j . Teniendo esta nueva notación en cuenta, la ecuación (1.28) anterior se puede reescribir como

$$\tilde{\Lambda}^{q_2}(s) = \mu_{\Lambda}^2(s) + \sum_{j=1}^{q_2} \left(\sum_{i=1}^{p_j} b_i^j \xi_i \right) g_j(s); \quad s \in (T_1, T_2) \quad (1.29)$$

Siendo este el modelo PCP $(q_2; p_1, \dots, p_j)$ para la predicción estocástica el proceso media.

Teniendo como base la predicción del proceso media, a partir de ella se puede obtener una predicción de los estadísticos de recuento que dependen de ella. En el capítulo [?] dedicado al PC compuesto, se aplicará la predicción obtenida del proceso media en la predicción de dichos estadísticos, ya que el PC compuesto es un caso más general y el PC no es no es más que una particularización suya.

1.4.2. Inferencia del proceso intensidad

En la sección anterior se ha hecho un estudio de la inferencia del PC en términos de su proceso media. Sin embargo, el PC puede expresarse también en función de su proceso intensidad. En Bouzas et al., 2006b [22] se realiza un primer intento de aplicación de ACPF al proceso intensidad de un PC. En dicho artículo, a partir de la observación de trayectorias muestrales del proceso de recuento, se aplicaba un estimador puntual para la intensidad en el contexto descrito en el mismo, y haciendo uso de la propiedad de independencia de los incrementos de un PC, se encontraban estimaciones de dicha intensidad en un conjunto discreto de instantes de tiempo para varias trayectorias. A estos datos, se les aplicaba el ACPF usual. Sin embargo, aunque el error de predicción obtenido era pequeño, en ningún momento llegaba a particularizarse el ACP, de modo que alguna de las propiedades teóricas del PC se conservase. El método descrito se mejora en Bouzas et al., 2010a [25] de la siguiente manera.

Una vez habiendo establecido una adaptación del ACPF a la estimación del proceso media de un PC conservando así su propiedad de monotonía, y sabiendo que la media, en caso de ser absolutamente continua, es la integral de la intensidad, se propone estimar la intensidad a partir de la derivación de la reconstrucción funcional monótona de las trayectorias de la media de la

ecuación (1.20) pertenecientes a \mathcal{C}^1 . Puesto que $\Lambda(t) = \int_{t_0}^t \lambda(\sigma) d\sigma$, se obtiene una estimación en \mathcal{C}^0 para las funciones muestrales de la intensidad diferenciando en la ecuación (1.20). Tomando un w fijo, el polinomio interpolado a trozos queda

$$p'_{wj} = \tilde{\Lambda}_w(t_j)H'_1(t) + \tilde{\Lambda}_w(t_{j+1})H'_2(t) + d_{wj}H'_3(t) + d_{wj+1}H'_4(t), \quad t \in [t_j, t_{j+1}]$$

$$j = 0, \dots, p-1 \quad (1.30)$$

No obstante, se obtiene de nuevo una reconstrucción funcional que no es apta para la aplicación del ACPF en su actual forma, necesita expresarse en todo el intervalo $[t_0, t_p]$. El polinomio que se ha obtenido es cuadrático y debía estar en \mathcal{C}^0 . Para ello, sería necesario expresar las funciones $H'_1(t)$, $H'_2(t)$, $H'_3(t)$ y $H'_4(t)$ en función de las funciones de la base cuadrática de Lagrange en \mathcal{C}^0 $\langle F_1, F_2, F_3 \rangle$ en el intervalo $[t_j, t_{j+1}]$.

$$aF_1 + bF_2 + cF_3 = \left(\frac{-4b + 2a + 2c}{h_j^2} \right) t^2 +$$

$$\left(\frac{-2at_{j+1} + (-t_j - t_{j+1}) + 4bt_j + 4bt_{j+1} + c(-t_j - t_{j+1}) - 2ct_j}{h_j^2} \right) t +$$

$$\frac{-at_{j+1}(-t_j - t_{j+1}) + ct_j(t_j + t_{j+1}) - 4bt_jt_{j+1}}{h_j^2} \quad (1.31)$$

Operando como se muestra en Bouzas et al., 2010a [25], las funciones derivadas se pueden escribir según las siguientes combinaciones lineales de las funciones de la base

$$H'_1(t) = -\frac{3}{2h_j}F_2(t), \quad H'_2(t) = \frac{3}{2h_j}F_2(t)$$

$$H'_3(t) = F_1(t) - \frac{1}{4}F_2(t), \quad F_3(t) - \frac{1}{4}F_2(t)$$

y así para cada $j = 0, \dots, p-1$, el polinomio a trozos que es estimación de la trayectoria muestral de intensidad es

$$p'_{wj}(t) = \hat{\Lambda}_w(t_j) \frac{-3}{2h_j} F_2(t) + \hat{\Lambda}_w(t_{j+1}) \frac{-3}{2h_j} F_2(t) \\ + d_{wj} \left(F_1(t) - \frac{1}{4} F_2(t) \right) + d_{wj+1} \left(F_3(t) - \frac{1}{4} F_2(t) \right), \quad t \in [t_j, t_{j+1})$$

donde

$$F_{1j} = \begin{cases} F_1, & t \in [t_j, t_{j+1}) \\ 0 & \text{en otro caso} \end{cases}, \quad F_{2j} = \begin{cases} F_2, & t \in [t_j, t_{j+1}) \\ 0 & \text{en otro caso} \end{cases} \\ \text{y } F_{3j} = \begin{cases} F_3, & t \in [t_j, t_{j+1}) \\ 0 & \text{en otro caso} \end{cases}$$

que puede reescribirse para dar una expresión para todo el intervalo $[t_0, t_p]$ como

$$p'_w(t) = \sum_{j=0}^p p'_{wj}(t), \quad t \in [t_0, t_p]$$

y reduciendo la expresión quedaría

$$p'_w(t) = \sum_{j=0}^{p-1} F_{2j}(t) \left[\frac{3}{2h_j} \left(\hat{\lambda}_w(t_{j+1}) - \hat{\Lambda}_w(t_j) \right) - \frac{d_{wj} + d_{wj+1}}{4} \right] \\ + \sum_{j=0}^{p-2} (F_{3j}(t) + F_{1j+1}(t)) d_{wj+1} + F_{10}(t) d_{w0} + F_{3p-1}(t) d_{wp} \quad (1.32)$$

Puede probarse fácilmente, que las funciones $F_{2j}(t)$, $j = 0, \dots, p-1$; $(F_{3j}(t) + F_{1j+1}(t))$, $j = 0, \dots, p-2$; $F_{10}(t)$ y $F_{3p-1}(t)$ son $2p+1$ funciones cuadráticas de \mathcal{C}^0 independientes y que forman una base en el intervalo completo $[t_0, t_p]$.

Una vez que se cuenta con las condiciones para poder aplicar el ACPF, y una vez aplicado este, se obtiene que el proceso intensidad se puede estimar mediante una descomposición truncada en función de las componentes principales

$$\lambda^q(t) = p'^q(t) = \mu_{p'}(t) + \sum_{j=1}^q \zeta_j f_j(t). \quad (1.33)$$

Este método de estimación se ilustra en Bouzas et al., 2010a [25], y prueba su mejora respecto al mencionado en Bouzas et al., 2006b [22]. Además en Bouzas et al., 2010b [26], se puede encontrar una aplicación práctica sobre datos reales referentes a la emisión de partículas radiactivas de dos isótopos distintos.

Una vez obtenida la estimación del proceso intensidad, esta podría utilizarse como base en los modelos PCP y predecir el PC, al igual que podría introducirse esa predicción en los estadísticos de recuento de modo análogo al caso de la media.

1.4.3. Contraste de bondad de ajuste

El clásico problema que aborda el contraste de bondad de ajuste es aquel que se centra en comprobar si ciertos datos provienen de una distribución de probabilidad dada. Sin embargo, a la hora de abordar si los datos vienen de un proceso estocástico el problema es mucho más complicado. Dado que las trayectorias muestrales de un proceso estocástico son funciones, lo normal pensar en el análisis de datos funcionales para la resolución del problema. Si buscamos en la literatura, encontramos que el contraste de bondad de ajuste habitualmente se realiza imponiendo una estructura prefijada o bien al proceso intensidad o a las trayectorias muestrales. Sin embargo, los datos muestrales disponibles de un proceso estocástico, suelen ser trayectorias muestrales observadas en un conjunto finito de instantes de tiempo. Si asumimos este hecho, el problema al que nos enfrentamos es averiguar si esas trayectorias muestrales provienen de un PC. Un test de hipótesis en el que la hipótesis nula es que el proceso de recuento observado es un cierto PC, y la hipótesis alternativa que no lo es, daría solución a dicho problema.

Es posible que el experimento sea conocido por experto del campo de

estudio del mismo y esto proporcione un modelo de PC, que sería el propuesto para la hipótesis nula del contraste, pero hay otras ocasiones en las que esto no sucede. Por esta razón se distinguen ambos casos, el contraste de bondad de ajuste dado un PC conocido y el contraste de bondad de ajuste a partir de los datos observados de un PC del que no se conoce su estructura estocástica.

Los contrastes de bondad de ajuste que se proponen en esta sección fueron inicialmente inspirados por el caso real del recuento de emisión de partículas por isótopos de Cesio 137 y Radio 226 con un contador de partículas de la Unidad de Radiofarmacia de la Universidad de Granada (ver Bouzas et al., 2010b [26]). Realmente, se construyó el contraste para resolver el problema real de distinguir si una nueva trayectoria observada después de un tiempo seguía el patrón inicial y por tanto, el contador estaba bien ajustado o si, por el contrario, no seguía el patrón y se podía concluir que el contador se había desajustado y necesitaba una puesta a punto. En el caso descrito, el proceso de recuento inicial que sirve de patrón, no era conocido y necesitaba estimarse previamente. Más adelante, se ha propuesto un contraste paralelo al anterior en el que el proceso de la hipótesis nula se considere conocido. Por orden de tipo más académico que de aparición en el tiempo, a continuación se describe el contraste con hipótesis nula conocida y luego el de cuya hipótesis que debe estimarse previamente.

Contraste de bondad de ajuste dado un PC

Abordemos en primer lugar el contraste de bondad de ajuste en el que se conoce la estructura estocástica del PC.

Sea $N(t)$ un PC con proceso intensidad $\lambda(t)$ determinado paramétricamente, por tanto la media y la varianza de la intensidad son conocidas y se denotan por $\mu(t)$ y $\sigma^2(t)$ respectivamente. Denotemos por $M(t)$ al proceso

de recuento que se observará para contrastar con $N(t)$.

Es un hecho que el trabajo computacional de los datos no siempre es práctico ni posible. Por ello, se propone utilizar un estimador de la intensidad para los datos observados que forman parte del contraste, ya que no requiere muchos tiempos de observación y que no conlleva cálculos complicados. Observando s trayectorias muestrales de $M(t)$ de tamaño $p \times (m + 1)$ en el subconjunto de puntos del tiempo t_j , $j = 0, \dots, p \times m$, se aplica el estimador puntual propuesto en Bouzas et al., 2006b [22] para cada una de las s trayectorias. El estimador puntual es

$$\hat{\lambda}_l(t_j) = \frac{1}{m} \sum_{i=1}^m \frac{M_i(t_j) - M_i(t_{j-1})}{t_j - t_{j-1}}, \quad j = 1, \dots, p; \quad l = 1, \dots, s \quad (1.34)$$

el cual es insesgado y consistente. Por tanto, para cada t fijo, se tienen s trayectorias muestrales estimadas de la intensidad en $j = 1, \dots, p$. El valor estimado de la intensidad en cada instante de tiempo t juega el papel de valor muestral para poder calcular el valor experimental de $\mu(t_j)$ denotado por $\mu_{exp}(t_j)$, $j = 1, \dots, p$. Así, si la intensidad es conocida y gaussiana o si $s \geq 30$, el intervalo de confianza para $\mu(t_j)$ es

$$I_j = \left[\mu_{exp}(t_j) \pm z_{1-\alpha/2} \sqrt{\frac{1}{s} \sigma^2(t_j)} \right], \quad j = 1, \dots, p \quad (1.35)$$

donde $z_{1-\alpha/2}$ es el percentil $(1 - \alpha/2)$ de una distribución Normal estándar.

Teniendo en cuenta I_j en el instante t_j , se puede concluir, siguiendo el razonamiento estadístico habitual que, tras calcular $\mu_{exp}(t_j)$, si $\mu_{exp}(t_j) \in I_j$ se acepta que la observación $M(t_j)$ es coherente con la propuesta $N(t_j)$, y que si por el contrario $\mu_{exp}(t_j) \notin I_j$ se rechaza. Lo cual equivale al siguiente contraste de hipótesis para un $\mu_{exp}(t_j) \notin I_j$ fijo

$$\begin{cases} H_0 : E[\lambda(t_j)] = \mu(t_j) \\ H_1 : E[\lambda(t_j)] \neq \mu(t_j) \end{cases} \quad (1.36)$$

donde $P[\mu_{exp}(t_j) \notin I_j \mid H_0 \text{ es cierta}] = \alpha$ es el error tipo I, es decir la probabilidad de rechazar la hipótesis nula siendo cierta y $P[\mu_{exp}(t_j) \in I_j \mid H_0 \text{ es falsa}] = \beta$ es el error tipo II, la probabilidad de aceptar la hipótesis nula siendo falsa. Como es habitual, se define el p-valor en el instante t_j como

$$p - \text{valor}(t_j) = 2P \left[Z > \left| \frac{\mu_{exp}(t_j) - \mu(t_j)}{\sqrt{\frac{1}{s}\sigma^2(t_j)}} \right| \right],$$

donde Z , como suele ser usualmente, es una variable aleatoria distribuida según una Normal estándar, y H_0 se acepta si $p - \text{valor}(t_j) > \alpha$. Siguiendo de nuevo el mismo razonamiento del contraste (1.36) para cada observación t_j , se obtiene el $p - \text{valor}(t_j)$, $j = 1, \dots, p$ para contrastar nuestro modelo. Como tenemos una trayectoria muestral experimental de $\mu(t)$ observada en p instantes de tiempo, no podemos aceptar o rechazar el modelo PC de la hipótesis nula basándonos tan solo en un punto del tiempo, habría que tener en cuenta toda la información disponible y utilizar la información de los p contrastes. Sobre la base de la Inferencia Simultánea, se propone utilizar el contraste de hipótesis múltiple propuesto por Benjamini y Hochber, 1995 [13] (procedimiento BH) para tomar la decisión. Este método trata de controlar la *False Discovery Rate* (FDR), es decir, la proporción esperada de falsos positivos en el conjunto de aquellos contrastes considerados significativos. Para ello a la hora de estimar el número de hipótesis nulas que son consideradas falsas no siéndolo, asume que todas las hipótesis nulas son ciertas, así la estimación de la FDR es mayor y por tanto conservadora. El propósito de seguir este procedimiento se basa en el hecho de que las variables $\lambda(t_j)$ son independientes para un PC y proporcionan resultados adecuados bajo estas condiciones. Considerando el siguiente contraste múltiple

$$\begin{cases} H_0^{(j)} : E[\lambda(t_j)] = \mu(t_j) \\ H_1^{(j)} : E[\lambda(t_j)] \neq \mu(t_j) \end{cases}, \quad j = 1, \dots, p \quad (1.37)$$

y los p-valores correspondientes mencionados con anterioridad $p - valor^{(1)} \leq \dots \leq p - valor^{(p)}$, H_0 se acepta si y solo si $p - valor^{(1)} > \alpha/p$, $p - valor^{(2)} > 2\alpha/p, \dots, p - valor^{(j)} > j\alpha/p, \dots, p - valor^{(p)} > \alpha$. Si al menos una de las comparaciones previas no fuera cierta, sea k el mayor i para el cual $p - valor^{(i)} \leq i\alpha/p$. Entonces, siguiendo el procedimiento BH, todas la hipótesis simples $H_0^{(j)}$ con $p - valor^{(j)}$ menor que $p - valor^{(k)}$ son rechazadas y por tanto en nuestro caso H_0 se rechaza también. Por tanto, se ha obtenido un contraste de bondad de ajuste para comprobar si los datos observados de un proceso de recuento son coherentes con un PC propuesto conocido su proceso de intensidad.

Contraste de bondad de ajuste dados los datos observados de un PC

Lo más habitual es que el PC no tenga una intensidad conocida paramétricamente. Por tanto, estudiaremos el caso en el que la información disponible sobre el PC, son las trayectorias muestrales observadas en un conjunto discreto de instantes de tiempo t_j , $j = 0, \dots, p$. Lo más recomendable es que los t_j sean los mismos instantes de tiempo que los observados para $M(t)$ con el fin de no complicar la implementación del programa. El estimador del proceso intensidad debe realizarse para poder conseguir un buen modelo, aunque es difícil. Por ello, proponemos estimar la media y la varianza con precisión para asumir que son cercanas a las reales. Habiendo estimado $\mu(t)$ y $\sigma^2(t)$, se denotan por $\hat{\mu}(t)$ y $\hat{\sigma}^2(t)$, haciendo unos ligeros cambios en el contraste de la sección anterior, el intervalo de confianza (1.35) será

$$\left[\mu_{exp}(t_j) \pm t_{1-\alpha/2, s-1} \sqrt{\frac{1}{s} \hat{\sigma}_j^2} \right]$$

donde $t_{1-\alpha/2, s-1}$ es el percentil $(1 - \alpha/2)$ de una distribución t de student. Entonces el contraste de hipótesis (1.36) pasa a ser

$$\begin{cases} H_0 : E[\lambda(t_j)] = \hat{\mu}(t_j) \\ H_1 : E[\lambda(t_j)] \neq \hat{\mu}(t_j) \end{cases}$$

y se tiene que

$$p - valor(t_j) = 2P \left[t > \left| \frac{\mu_{exp}(t_j) - \hat{\mu}(t_j)}{\sqrt{\frac{1}{s} \hat{\sigma}_j^2}} \right| \right]$$

donde t es una variable distribuida según una t de student con $s - 1$ grados de libertad. Por último, el contraste de bondad de ajuste para contrastar si las trayectorias muestrales observadas de $M(t)$ son coherentes con las observadas de $N(t)$ es

$$\begin{cases} H_0^{(j)} : E[\lambda(t_j)] = \hat{\mu}(t_j) \\ H_1^{(j)} : E[\lambda(t_j)] \neq \hat{\mu}(t_j) \end{cases}, \quad j = 1, \dots, p$$

y como se ha explicado anteriormente, H_0 se acepta si y solo si $p - valor^{(1)} > \alpha/p$, $p - valor^{(2)} > 2\alpha/p, \dots, p - valor^{(j)} > j\alpha/p, \dots, p - valor^{(p)} > \alpha$, en otro caso se rechaza. Obteniéndose así el contraste de bondad de ajuste en el caso que la estructura estocástica del PC es desconocida, es decir, a partir de los datos observados del proceso.

Capítulo 2

Proceso de Cox compuesto

Un proceso puntual marcado es un proceso puntual en el cual cada ocurrencia tiene asociada una variable aleatoria. Snyder y Miller, 1991 [117] definieron el *proceso de Poisson compuesto* como un proceso puntual marcado tal que verifica que el proceso puntual es de Poisson y que las marcas de las ocurrencias, son variables aleatorias idénticamente distribuidas y a su vez independientes del proceso de Poisson. Muchos campos de estudio utilizan el PC compuesto como por ejemplo en geología, donde se utilizan a la hora de modelizar fenómenos reales, un clásico ejemplo es el de los terremotos registrados, donde son interpretados como un PC compuesto en el que las magnitudes serían las marcas del proceso, como puede verse en Ogata, 1998 [92], Gospodinov, 2001 [58]. Ejemplos en demografía, econometría, teoría de riesgos, astrofísica, entre otros pueden encontrarse en Si, 2001 [114], Economou, 2003 [47], Borak et al., 2005 [16], Lin y Pavlova, 2006 [80], Chertok et al, 2016 [37], entre otros.

Este capítulo se centra en el estudio del PC compuesto, que es una ampliación del proceso de Poisson compuesto, que es un proceso puntual marcado que tiene como base el PC.

2.1. Definición

Como ya se ha adelantado, el PC compuesto, es, en pocas palabras, la extensión natural del proceso de Poisson compuesto, en el cual la intensidad es un proceso estocástico influenciado por un proceso externo, el proceso información. La definición formal del PC compuesto es la siguiente:

Definición 2.1.1 *Se define el PC compuesto, como el proceso de recuento con proceso intensidad $\{\lambda(t, x(t)) : t \geq t_0\}$ donde $\{x(t) : t \geq t_0\}$ es el proceso información y con marcas i.i.d. asociadas a cada suceso. El n -ésimo tiempo de ocurrencia se notará como w_n y su correspondiente marca como u_n , que es una realización de la variable aleatoria U . Siendo su acumulador de marcas*

$$z(t) = \sum_{i=0}^{N(t)} u_i$$

donde $u_0 = 0$.

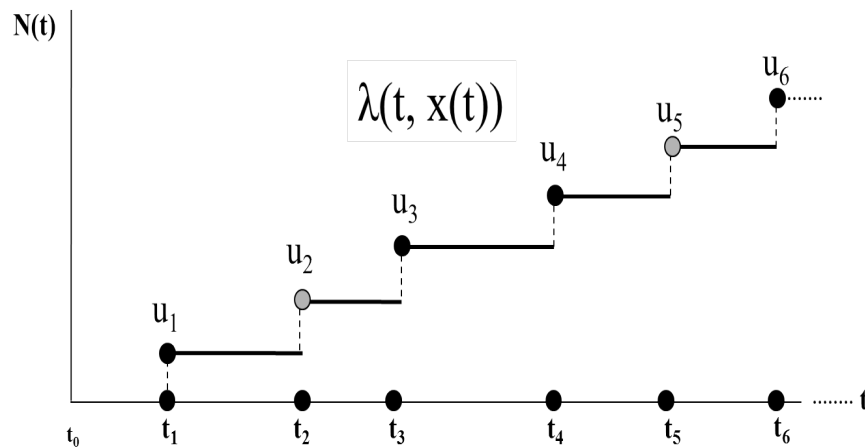


Figura 2.1: Proceso de Cox compuesto con marcas asociadas.

El PC compuesto es un modelo más flexible, gracias a la aleatoriedad de su intensidad, que ha sido objeto de estudio de distintos autores entre los que

cabe destacar Brèmaud, 1981 [34], Daley y Vere-Jones, 1988 [41], Snyder y Miller, 1991 [117], Last y Brandt, 1995 [76], Møller y Waagepetersen, 2004 [86], Bouzas et al., 2004 [20], centrándose algunos de ellos en el estudio de estos procesos desde el punto de vista de las martingalas. Aunque es cierto que en pocas ocasiones se utilizan en ejemplos con datos reales, si se pueden encontrar algunas aplicaciones en Brix, 2002 [35], García et al., 2004 [54] y Barta et al., 2005 [10]. Puede que el principal motivo por el cual no suele aplicarse a ejemplos con datos reales, sea la dificultad de estimación que tienen los parámetros media o intensidad del PC, así como su caracterización.

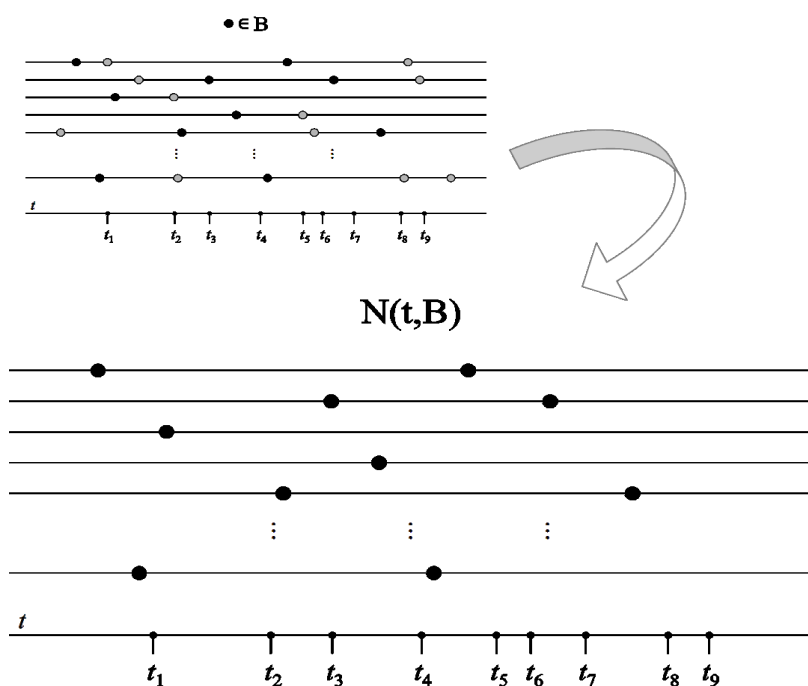


Figura 2.2: Proceso de Cox compuesto con marcas en un subconjunto B .

De igual modo que hace el proceso de Poisson compuesto, el PC compuesto produce un sencillo proceso acumulador de marcas $z(t)$, que es la suma de todas las marcas hasta un punto dado en el tiempo, como se ha visto en la definición (2.1.1). Aunque no es objetivo de esta memoria, merece la pena

al menos mencionarse que es posible extender el PC compuesto al PC compuesto filtrado, proceso que modeliza la respuesta de los puntos del proceso puntual marcado. En Bouzas et al., 2004 [20] se desarrolla un estudio de este proceso.

2.2. Teoremas de representación

No es usual el estudio del PC compuesto con marcas en un subconjunto dado de marcas específicas. Considerando las marcas en los puntos del proceso de recuento, puede que sean interesantes solo aquellos puntos con unas marcas concretas según su significado. Es decir, puede que nos interese estudiar un proceso puntual extraído del PC compuesto original. Entonces, es pertinente preguntarse acerca de la naturaleza de este otro proceso puntual. Los teoremas de representación abordan y resuelven este problema.

Teniendo como base los teoremas de representación del proceso de Poisson compuesto y el método de condicionamiento de Snyder y Miller, 1991 [117], cuyas demostraciones se basan en el funcional característico del proceso, es posible dar una versión extendida al caso doblemente estocástico, como puede verse en Bouzas et al., 2007 [24], se exponen los teoremas de representación del PC compuesto. Estos teoremas a parte de la importancia de su interpretación, facilitan el cálculo de estadísticos de proceso.

Dado el acumulador de marcas $z(t)$, y utilizando el método de condicionamiento de Snyder y Miller, 1991 [117], el funcional característico del proceso puede expresarse como

$$\Phi_z = E_x \left[\exp \left\{ \int_{t_0}^t \lambda(t, x(t)) (M_u(jv(t)) - 1) dt \right\} \right], \quad (2.1)$$

donde la esperanza es condicional al proceso información, $j = \sqrt{-1}$, $\{v(t); t_0 \leq t < T\}$ es una función arbitraria vectorial y M_u es la función característica de U , variable aleatoria asociada a las marcas.

Seguidamente se enuncian los citados teoremas de representación expuestos en Bouzas et al., 2007 [24], haciendo distinción tanto para el caso con espacio de marcas numerable, como para el no numerable

Teorema 2.2.1 Teorema de representación. Caso: Espacio de marcas numerable

Se considera un PC compuesto cuyas marcas pertenecen a un espacio numerable $\{U_1, U_2, \dots, U_k, \dots\}$ donde la marca U_k ocurre en un punto con probabilidad p_k , independientemente de la ocurrencia en otros puntos. Se define $\{N(t, U_k); t \geq t_0\}$ como el proceso de recuento que cuenta los puntos con marca U_k . Dado que las marcas son independientes, esto implica que los procesos de recuento $\{N(t, U_k); t \geq t_0\}$ con $k = 1, 2, \dots$ son mutuamente independientes y son a su vez PC con intensidades $\{p_k \lambda(t, x(t)); t \geq t_0\}$ con $k = 1, 2, \dots$. En este caso, el acumulador de marcas $z(t)$ puede expresarse como

$$z(t) = \sum_{k=0}^{\infty} U_k N(t, U_k), \quad t \geq t_0.$$

Teorema 2.2.2 Teorema de representación. Caso: Espacio de marcas no numerable

Se considera un PC compuesto cuyas marcas pertenecen a un espacio \mathcal{U} no numerable con función de densidad $P_u(U)$. Se define $\{N(t, B_k); t \geq t_0\}$ como el proceso de recuento que cuenta los puntos que tienen marca en el subconjunto $B_k \subset \mathcal{U}$. Tomando los B_k , $k = 1, 2, \dots$ como regiones disjuntas del espacio \mathcal{U} , los procesos de recuento $\{N(t, B_k); t \geq t_0\}$ con $k =$

$1, 2, \dots$ son mutuamente independientes y son a su vez PC con intensidades $\left\{ \lambda(t, x(t)) \int_{B_k} P_u(U); t \geq t_0 \right\}$ con $k = 1, 2, \dots$. Se representa el acumulador de marcas $z(t)$ como

$$z(t) = \int_{\mathcal{U}} U N(t, dU),$$

donde el término $\{N(t, dU); t \geq t_0\}$ se interpreta como el PC que cuenta los puntos en el intervalo $[t_0, t)$ que tienen marcas en $[U, U + dU)$.

Resumiendo, si no se pretende estudiar todos los puntos, sino solo aquellos que tienen marcas específicas, es posible, siempre y cuando se tengan en cuenta los procesos puntuales con marcas en el subconjunto adecuado del espacio de marcas, y se trate como el PC que es. Por tanto, no es necesario extraer el proceso $N(t, B_k)$ del PC compuesto completo y original, ni tampoco estudiar su proceso intensidad o proceso media como un PC compuesto nuevo y diferente, sino que se puede utilizar el estudio de su proceso intensidad o su proceso media mediante las intensidades dadas en los teoremas de representación. Este hecho tiene especial relevancia cuando los procesos intensidad o media han sido estimados, puesto que no habría que repetir la estimación del proceso para cada subconjunto B_k en el que se esté interesado, incluso para resolver los problemas de marcas perdidas, entre otros casos. Un ejemplo puede verse en Bouzas et al., 2007 [24] donde se presentan simulaciones para las cuales se estima la media del PC mediante Análisis Funcional en Componentes Principales y esta a su vez, se utiliza para la estimación en distintos subconjuntos B_k . Desde el punto de vista práctico, un ejemplo sería teniendo sismos registrados a lo largo del tiempo, considerar solo aquellos en los que su magnitud ha superado un límite dado, o cuyo epicentro no es tan profundo como el establecido o que se encuentre en un área determinada.

Por otro lado, es interesante señalar que el PC es un caso particular del

PC compuesto, en el que las marcas están en un espacio numerable con marca $U = 1$. Sin embargo, si se escoge el espacio de marcas no numerable $\mathbb{R} = (-\infty, +\infty)$, el PC compuesto puede representar un PC con saltos no unitarios, lo que significa que la propiedad de regularidad del proceso de Poisson se ha relajado. Otro claro ejemplo es el PC compuesto con supresión aleatoria, ya que puede verse como un PC compuesto con marcas en el espacio $\{0, 1\}$ donde 0 es la marca de las ocurrencias borradas.

2.3. Estadísticos de recuento de un PC compuesto con marcas específicas

Siguiendo el esquema del capítulo anterior, esta sección se centrará en el estudio de los estadísticos de recuento del proceso. Para ello, en un primer lugar se muestra el proceso media del PC compuesto con marcas en un subconjunto del espacio de marcas, para así poder hacer uso de él en aquellos otros estadísticos que tienen su expresión en función de dicho proceso.

Con el fin de no complicar la notación los estadísticos serán dados para un subconjunto cualquiera B . Por tanto, el PC compuesto con marcas en B , se notará en adelante por $N(t, B)$.

2.3.1. Proceso media

Los teoremas de representación pueden ser expresados en función del proceso media. Como es sabido, en un PC cuya media es absolutamente continua, esta es la integral del proceso intensidad. Este resultado implica que los PC con marcas en subconjuntos disjuntos B_k que son a su vez mutuamente independientes, tienen como procesos media:

- **Caso numerable**

$$p_k \int_{t_0}^t \lambda(\sigma, x(\sigma)) d\sigma = p_k \Lambda(t, x(t)).$$

- **Caso no numerable**

$$\int_{t_0}^t \int_B \lambda(\sigma, x(\sigma)) P_u(dU) = \Lambda(t, x(t)) \int_B P_u(dU), \quad (2.2)$$

donde se asume que $N(t_0) = 0$ *c.s.*, lo cual no es restrictivo.

Una vez establecidas las expresiones para el proceso media de un PC con marcas específicas, se puede hacer uso de ellas en los estadísticos de recuento básicos de $N(t, B)$, y dar su fórmula tanto en términos del proceso intensidad como de la media, para ello tendremos en cuenta el caso con espacio de marcas no numerable (el caso numerable es análogo).

2.3.2. Función masa de probabilidad

La función masa de probabilidad de un PC compuesto con marcas específicas, puede expresarse en función del proceso media como sigue

$$\begin{aligned} P[N(t, B) = n] &= \\ E_x \left[\frac{1}{n!} \left(\int_{t_0}^t \int_B \lambda(\sigma, x(\sigma)) P_u(dU) d\sigma \right)^n \exp \left\{ - \int_{t_0}^t \int_B \lambda(\sigma, x(\sigma)) P_u(dU) d\sigma \right\} \right] \\ &= E_x \left[\frac{1}{n!} \left(\Lambda(\sigma, x(\sigma)) \int_B P_u(dU) \right)^n \exp \left\{ - \Lambda(\sigma, x(\sigma)) \int_B P_u(dU) \right\} \right] \end{aligned}$$

Y en términos de la función generatriz de momentos toma la siguiente forma $\Lambda(t, x(t)) \int_B P_u(dU)$, $G_{\Lambda(t, x(t)) \int_B P_u(dU)}$ como

$$P[N(t, B) = n] = \frac{1}{n!} G_{\Lambda(t, x(t)) \int_B P_u(dU)}(-1) \quad (2.3)$$

De igual modo, puede darse una expresión para la media del proceso $N(t, B)$

$$\begin{aligned} E[N(t, B)] &= \\ E_x \left[\int_{t_0}^t \int_B \lambda(\sigma, x(\sigma)) d\sigma P_u(dU) \right] &= E_x \left[\Lambda(t, x(t)) \int_B P_u(dU) \right] \end{aligned} \quad (2.4)$$

Como ya se ha indicado anteriormente, la expresión de la función masa de probabilidad del proceso $N(t, B)$ es de difícil tratamiento en la mayoría de los casos, haciendo que sea complicado su cálculo. Sin embargo, En Bouzas et al., 2007 [24] se puede encontrar un ejemplo en el que se calcula la función masa de probabilidad de un PC compuesto con intensidad de banda estrecha con marcas específicas, para ello se hizo uso de una fórmula recursiva.

2.3.3. Moda

La moda del proceso $N(t, B)$ es su valor más probable y se notará en adelante como $n_{m\acute{a}x} \in \mathbb{N}$. Para este estadístico puede hallarse un intervalo dentro del cual se conoce que se encuentra el valor con máxima probabilidad. Su estudio completo puede encontrarse en Bouzas et al., 2011 [28]. Hemos de notar que tomando B como todo el espacio de marcas, o simplemente eliminando el término que depende de B , $\int_B P_u(dU)$, se lograría una demostración análoga a la que se expone a continuación para el caso del PC, tal y como se avanzó en la Sección (1.2.4) del capítulo anterior.

Proposición 2.3.1 *Sea el proceso $N(t, B)$ un PC compuesto con marcas en $B \subset \mathcal{U}$. El número de ocurrencias con mayor probabilidad del proceso, $n_{m\acute{a}x}$, en un instante fijo es $\lceil \Lambda(t, x(t)) \int_B P_u(dU) - 1 \rceil$ si esta expresión pertenece a \mathbb{N} , e $\text{int} \lceil \Lambda(t, x(t)) \int_B P_u(dU) - 1 \rceil$ o $\text{int} \lceil \Lambda(t, x(t)) \int_B P_u(dU) - 1 \rceil + 1$ si no pertenece a \mathbb{N} y será 0 si la expresión es negativa.*

Nota: $\text{int} =$ parte entera.

Demostración. Por definición, la moda $n_{m\acute{a}x}$, es el entero n tal que se alcanza $\acute{m}ax P[N(t, B) = n]$, donde la probabilidad $P[N(t, B) = n]$ es la función masa de probabilidad de un PC compuesto con marcas específicas dada anteriormente.

Para un t fijo, se conoce el valor de $\Lambda(t, x(t))$ dado el proceso información, y entonces

$$\begin{aligned} \text{máx} P[N(t, B) = n] = \\ \text{máx} \left[\frac{1}{n!} \left(\Lambda(t, x(t)) \int_B P_u(dU) \right)^n \exp \left[-\Lambda(t, x(t)) \int_B P_u(dU) \right] \right] \end{aligned}$$

Puesto que $\exp \left[-\Lambda(t, x(t)) \int_B P_u(dU) \right]$ es constante, tan solo habría que maximizar la expresión $\frac{1}{n!} \left(\Lambda(t, x(t)) \int_B P_u(dU) \right)^n$. Para simplificar la notación, llamaremos $A = \Lambda(t, x(t)) \int_B P_u(dU)$. Se verifica que:

1. Para un $A \in \mathbb{R}$ conocido, $\frac{A^n}{n!} \xrightarrow{n \rightarrow \infty} 0$ luego existe $n_0 \in \mathbb{N}$ a partir del cual la serie decrece en n . Entonces,

$$\frac{A^n}{n!} \geq \frac{A^{n+1}}{(n+1)!} \quad \text{si y solo si } n \geq A - 1$$

2. Además,

$$\frac{A^n}{n!} \leq \frac{A^{n+1}}{(n+1)!} \quad \text{si y solo si } n \leq A - 1$$

de este modo, en el caso que $A \in \mathbb{N}$, $n_{\text{máx}} = A - 1$ y si $A \notin \mathbb{N}$, entonces $n_{\text{máx}} = \text{int}(A - 1)$ o $n_{\text{máx}} = \text{int}(A - 1) + 1$. Hay que tener en cuenta la posibilidad de que resulte un valor negativo, lo que ocurriría si $A < 1$, entonces se tomaría $n_{\text{máx}} = 0$. ■

Para que resulte más cómodo, puede escribirse la moda $n_{\text{máx}}$ como

$$n_{\text{máx}}^B(t) = \begin{cases} \Lambda(t) \int_B P_u(dU) - 1, & \Lambda(t) \int_B P_u(dU) \in \mathbb{N} \\ \begin{cases} \text{int}(\Lambda(t) \int_B P_u(dU) - 1) \\ \text{or} \\ \text{int}(\Lambda(t) \int_B P_u(dU) - 1) + 1 \end{cases}, & \Lambda(t) \int_B P_u(dU) \notin \mathbb{N} \\ 0, & \Lambda(t) \int_B P_u(dU) < 1 \end{cases} \quad (2.5)$$

2.3.4. Función de distribución acumulada

Se define la función de distribución acumulada del proceso $N(t, B)$ como la probabilidad de haber observado menos de un número dado de ocurrencias en B antes de un instante de tiempo fijo. Dada la función masa de probabilidad, la expresión de la función de distribución acumulada es la siguiente

$$\begin{aligned}
 P[N(t, B) < n] &= P[N(t, B) \leq n - 1] = \\
 E_x \left[\sum_{i=0}^{n-1} \frac{1}{i!} \left(\Lambda(t) \int_B P_u(dU) \right)^i \exp \left[-\Lambda(t) \int_B P_u(dU) \right] \right] &= \\
 E_x \left[\frac{\Gamma(n, \Lambda(t) \int_B P_u(dU))}{\Gamma(n)} \right] & \quad (2.6)
 \end{aligned}$$

donde $\Gamma(n) = \int_0^\infty \sigma^{n-1} e^{-\sigma} d\sigma$ es la conocida función gamma y, a su vez, $\Gamma(n, x) = \int_x^\infty \sigma^{n-1} e^{-\sigma} d\sigma$ es la función gamma incompleta.

2.4. Estadísticos de tiempo de un PC compuesto con marcas específicas

Una vez expuesto los estadísticos de recuento, pasamos en esta sección al estudio de los estadísticos de los tiempos de ocurrencia de un PC compuesto con marcas específicas en un subconjunto B .

2.4.1. Función de supervivencia

La función de supervivencia, también conocida como función de fiabilidad, de un PC compuesto $N(t, B)$ es la probabilidad de que el n -ésimo punto del proceso suceda después de un tiempo T . Esta función está relacionada con la función de distribución acumulada del siguiente modo

$$P^B[W_n > T] = P[N(t, B) < n] \quad (2.7)$$

donde P^B indica la probabilidad dentro del subconjunto de marcas B y por tanto su expresión es la misma que la de la ecuación (2.6).

Haciendo uso una vez más del método de condicionamiento de Snyder y Miller, 1991 [117], se obtiene que la *función de densidad condicionada* de los n primeros tiempos de ocurrencia de $N(t, B)$ es

$$\begin{aligned}
& P_{w_n/w_{n-1}}^B(W_n/W_{n-1}) \\
&= E_x \left[\lambda(W_n) \int_B P_u(dU) \exp \left[- \int_{W_{n-1}}^{W_n} \lambda(\sigma) P_u(dU) d\sigma \right] \right] \\
&= E_x \left[\lambda(W_n) \int_B P_u(dU) \exp \left[-(\Lambda(W_n) - \Lambda(W_{n-1})) \int_B P_u(dU) \right] \right] \quad (2.8)
\end{aligned}$$

2.4.2. Probabilidad de tener una nueva ocurrencia en un intervalo de tiempo

Otro estadístico de gran interés es aquel que proporciona la probabilidad de que la siguiente ocurrencia con marca B esté en el intervalo de tiempo $[W_{n-1}, T)$, cuya expresión es la siguiente

$$\begin{aligned}
& P_{w_n/w_{n-1}}^B(W_n < T/W_{n-1}) \\
&= E_x \left[\int_{W_{n-1}}^T \lambda(W_n) \int_B P_u(dU) \exp \left[- \int_{W_{n-1}}^{W_n} \int_B \lambda(\sigma) P_u(dU) d\sigma \right] dW_n \right] \\
&= E_x \left[-\exp \left[- \int_{W_{n-1}}^T \int_B \lambda(\sigma) P_u(dU) d\sigma \right] + \exp \left[- \int_{W_n}^T \int_B \lambda(\sigma) P_u(dU) d\sigma \right] \right] \\
&= E_x \left[1 - \exp \left[- \int_{W_{n-1}}^T \int_B \lambda(\sigma) P_u(dU) d\sigma \right] \right] \\
&= E_x \left[1 - \exp \left[-(\Lambda(T) - \Lambda(W_{n-1})) \int_B P_u(dU) \right] \right] \quad (2.9)
\end{aligned}$$

2.4.3. Función de supervivencia condicionada

Como consecuencia del estadístico anterior, se puede escribir la función supervivencia o fiabilidad condicionada del siguiente modo

$$P_{w_n/w_{n-1}}^B(W_n > T/W_{n-1}) = E_x \left[\exp \left[-(\Lambda(T) - \Lambda(W_{n-1})) \int_B P_u(dU) \right] \right] \quad (2.10)$$

Teniendo en cuenta la relación entre los tiempos de ocurrencia y los tiempos de llegadas, $w_n = w_{n-1} + t_n$, donde t_n es el tiempo de llegada n -ésimo, se puede llegar a la expresión de la densidad de ocurrencia adelantada.

2.4.4. Densidad de ocurrencia adelantada

La densidad de la ocurrencia adelantada es igual a la densidad condicionada de t_n y se expresa como

$$p_{t_n/w_{n-1}}^B(T/W_{n-1}) = p_{w_n/w_{n-1}}^B(W_n + T/W_{n-1}) = E_x \left[\lambda(W_n + T) \int_B P_u(dU) \exp \left[-(\Lambda(W_n + T) - \Lambda(W_{n-1})) \int_B P_u(dU) \right] \right]$$

Es de señalar, que una característica común a los estadísticos de recuento y a los estadísticos de tiempo de un PC compuesto con marcas específicas, es que todos ellos están expresados en función de la media del PC respecto a las marcas, lo que permitirá utilizar cualquier estudio adicional de dicha media para todo proceso $N(t, B)$.

2.5. Inferencia

Para poder abordar la estimación de los estadísticos de un PC compuesto con marcas específicas, se requiere el uso de varios resultados ya expuestos a lo largo de esta memoria. Por un lado, se puede estimar la media del proceso sin tener en cuenta sus marcas $N(t)$, como ya se presentó en la Sección 1.4.1, del capítulo anterior. Por lo que, a partir de su predicción vista en la ecuación (1.29) y haciendo uso de la independencia entre las marcas y ocurrencias del PC compuesto, se pueden predecir varios de sus estadísticos de recuento. Por otra parte, se debe recordar que la media del proceso $N(t, B)$ se expresa en términos de la media de $N(t)$ y una constante dependiente de B según los teoremas de representación. Por tanto, surge una nueva aplicación de la estimación del proceso media del PC base del PC compuesto, la estimación de los estadísticos de $N(t, B)$.

2.5.1. Inferencia de los estadísticos de recuento

En este apartado haciendo uso tanto de los teoremas de representación como de la predicción del proceso media, se darán expresiones para los estadísticos de recuento de un PC compuesto en un tiempo futuro $s \in (T_1, T_2)$, habiendo tomado la observación del proceso en el intervalo $[T_0, T_1]$, como se muestra en Bouzas et al., 2010 [27]. Esto es, si una trayectoria muestral ha sido observada hasta un tiempo T_1 y se necesita una predicción en s , debido a la independencia de los incrementos del PC compuesto cabe la posibilidad de estimar $N(s, B)$ en (T_1, T_2) y sumarle las ocurrencias en el pasado. De este modo, se obtienen los estimadores de los principales estadísticos de recuento que se muestran a continuación.

Media

La media del proceso $N(T_1, s, B)$, $s \in (T_1, T_2]$, donde $N(T_1, s, B)$ es el número de puntos entre T_1 y s , viene dada por

$$E [N(T_1, s, B)] \simeq E_x \left[\Lambda(s) \int_B P_u(dU) \right] = \mu_\Lambda^2(s) \int_B P_u(dU) \quad (2.11)$$

Moda

La moda de $N(T_1, s, B)$, $s \in (T_1, T_2)$ se define como el número de ocurrencias con probabilidad máxima hasta el instante s , $n_{\text{máx}}^B(s)$, siendo su expresión

$$n_{\text{máx}}^B(s) \simeq \quad (2.12)$$

$$\begin{cases} \tilde{\Lambda}^{q_2}(s) \int_B P_u(dU) - 1, & \tilde{\Lambda}^{q_2}(s) \int_B P_u(dU) \in \mathbb{N} \\ \left\{ \begin{array}{l} \text{int} \left(\tilde{\Lambda}^{q_2}(s) \int_B P_u(dU) - 1 \right) \\ \text{or} \\ \text{int} \left(\tilde{\Lambda}^{q_2}(s) \int_B P_u(dU) - 1 \right) + 1 \end{array} \right. & , \tilde{\Lambda}^{q_2}(s) \int_B P_u(dU) \notin \mathbb{N} \\ 0, & \tilde{\Lambda}^{q_2}(s) \int_B P_u(dU) < 1 \end{cases} \quad (2.13)$$

Función de distribución acumulada

La función de distribución acumulada de $N(t, B)$ en $s \in (T_1, T_2)$ es la probabilidad de haber observado menos de un cierto número de ocurrencias en B hasta el instante de tiempo s . Haciendo uso de las ecuaciones de la función de distribución acumulada de un PC (2.6), de la predicción de la media (1.29) y notando $N(T_1, B) = n^{T_1}$, se obtiene

$$P [N(s, B) < n] \simeq \begin{cases} 0, & n \leq n^{T_1} \\ E_x \left[\frac{\Gamma \left(n - n^{T_1} - 1, \Lambda(s) \int_B P_u(dU) \right)}{\Gamma(n - n^{T_1} - 1)} \right], & n > n^{T_1} \end{cases}$$

donde $\Gamma(n)$ es la función gamma y $\Gamma(n, x)$ es la función gamma incompleta.

2.5.2. Inferencia de los estadísticos de tiempo

Pasemos ahora a los estadísticos de los tiempos de ocurrencia. Supongamos que se ha observado una trayectoria muestral en un conjunto discreto de instantes de tiempo de un intervalo $[T_0, T_1]$ y se necesita una estimación en s . Al igual que en los estadísticos de recuento, teniendo en cuenta la independencia de los incrementos del PC compuesto y las ecuaciones de los estadísticos que se mostraron en la sección (2.3), se puede estimar el proceso $N(s, B)$ en el intervalo (T_1, T_2) , conocido en un intervalo previo dando lugar a la ecuación de la media (2.11).

Sin embargo, la aplicación de la estimación de la media (1.29) da lugar a una expresión intratable de los estadísticos de tiempo de un PC compuesto con marcas específicas, aunque es habitual utilizarla con el PC. No obstante, a pesar de esto, si se verifican unas restricciones muy débiles, es posible derivar una expresión que se puede implementar, como se muestra en la siguiente proposición.

Proposición 2.5.1 *Dada la estimación del proceso media de $N(t)$ de la ecuación (1.29), se asume que las componentes principales ξ_i utilizadas para la estimación de las componentes del futuro η_j , $j = 1, \dots, q_2$ son independientes y se distribuyen según una distribución Normal (o $k > 30$) y sus distribuciones son (o pueden estimarse como) $\mathcal{N}(0, \sigma_i)$ donde σ_i^2 son las varianzas estimadas previamente en el correspondiente ACPF.*

Para un tiempo fijo $s \in (W_{n-1} = T_1, T_2)$, la probabilidad de haber una

nueva ocurrencia con marca en B en el intervalo ($W_{n-1} = T_1, s$) es

$$\begin{aligned}
 P_{w/w_{n-1}}^B(W_n < s/W_{n-1}) &\simeq \\
 &1 - \exp \left[\left(\Lambda^{q_1}(W_{n-1}) - \mu_{\Lambda}^2(s) \right) \int_B P_u(dU) \right] \\
 &\cdot \exp \left[\frac{(\int_B P_u(dU))^2}{2} \sum_{j=1}^{q_2} \left(\sum_{i=1}^{p_j} (\sigma_i b_i^j)^2 \right) (g_j(s))^2 \right] \quad (2.14)
 \end{aligned}$$

donde p es el mayor p_j , $\forall j$

Demostración. Con el fin de no complicar la notación se denotará por $a = \int_B P_u(dU)$ y $b = \Lambda^{q_1}(W_{n-1})$, ambas constantes. Entonces la ecuación (2.9) puede reescribirse como

$$\begin{aligned}
 P_{w/w_{n-1}}^B(W_n < s/W_{n-1}) &= E_x \left[1 - \exp \left[-(\Lambda(s) - \Lambda(W_{n-1})) \int_B P_u(dU) \right] \right] \\
 &= E_x [1 - \exp [-a(\Lambda(s) - b)]] \\
 &= 1 - E_x [e^{-a\Lambda(s)} e^{ab}] = 1 - e^{ab} M_{\Lambda(s)}(-a)
 \end{aligned}$$

donde $M_{\Lambda(s)}(\sigma)$ es la función generatriz de $\Lambda(s)$. Utilizando la ecuación de la predicción de la media (1.29), se obtiene

$$\begin{aligned}
 P_{w/w_{n-1}}^B(W_n < s/W_{n-1}) &\simeq 1 - e^{ab} E \left[\exp \left[-a \left(\mu_{\Lambda}^2(s) + \sum_{j=1}^{q_2} \left(\sum_{i=1}^{p_j} b_i^j \xi_i \right) g_j(s) \right) \right] \right] \\
 &\quad (2.15)
 \end{aligned}$$

Sustituyendo a y b por su expresiones se obtendría la probabilidad, pero como ya se comentó es una expresión intratable, difícil de calcular. Con el fin de obtener una expresión práctica para esta probabilidad, introducimos la

siguiente notación. Llamaremos p al mayor p_j , $\forall j$. Es decir, p es el mayor subíndice de las componentes principales del pasado utilizadas en la estimación de las componentes principales del futuro. Por tanto la ecuación (2.15) puede reescribirse como

$$\begin{aligned}
&\simeq 1 - e^{ab} E \left[\exp \left[-a \left(\mu_{\Lambda}^{(2)}(s) + \sum_{j=1}^{q_2} \left(\sum_{i=1}^p b_i^j \xi_i \right) g_j(s) \right) \right] \right] \\
&= 1 - e^{ab} e^{-a\mu_{\Lambda}^{(2)}(s)} E \left[\exp \left[-a \sum_{j=1}^{q_2} \left(\sum_{i=1}^p b_i^j \xi_i \right) g_j(s) \right] \right] \\
&= 1 - e^{ab} e^{-a\mu_{\Lambda}^{(2)}(s)} E \left[\exp \left[\sum_{i=1}^p \xi_i \left(\sum_{j=1}^{q_2} -a b_i^j g_j(s) \right) \right] \right] \\
&= 1 - e^{ab} e^{-a\mu_{\Lambda}^{(2)}(s)} E \left[\prod_{i=1}^p \prod_{j=1}^{q_2} \exp \left[-a b_i^j \xi_i g_j(s) \right] \right]
\end{aligned}$$

Debe notarse que algunas componentes principales ξ_i pueden faltar debido a que $b_i^j = 0$, $\forall j$, lo que significa que no aparecen en la estimación de las componentes principales del futuro. Si el resto de las componentes principales ξ_i son independientes, la expresión de la ecuación (2.16) puede expresarse como

$$\begin{aligned}
&P_{w/w_{n-1}}^B(W_n < s/W_{n-1}) \\
&\simeq 1 - e^{ab} e^{-a\mu_{\Lambda}^{(2)}(s)} \left[\prod_{i=1}^p \prod_{j=1}^{q_2} M_{\xi_i}(-a b_i^j g_j(s)) \right]
\end{aligned}$$

Por último, se asume que las componentes principales ξ_i , se distribuyen mediante $\mathcal{N}(0, \sigma_i^2)$ donde σ_i^2 son estimadas en el ACPF. En el caso que esto no se verifique pero el tamaño muestral sea mayor que 30, se pueden aproximar sus distribuciones mediante Normales. Entonces, como la función generatriz

de ξ_i es $M_{\xi_i}(u) = \exp\left[\frac{\sigma_i^2 u^2}{2}\right]$, la ecuación se puede reescribir

$$\begin{aligned}
& P_{w/w_{n-1}}^B(W_n < s/W_{n-1}) \\
& \simeq 1 - \exp\left[\left(\Lambda(W_{n-1}) - \mu_\Lambda^2(s)\right) \int_B P_u(dU)\right] \\
& \times \left[\prod_{i=1}^p \prod_{j=1}^{q_2} \exp\left[\frac{(\sigma_i a b_i^j g_j(s))^2}{2}\right]\right] \\
& = 1 - \exp\left[\left(\Lambda(W_{n-1}) - \mu_\Lambda^2(s)\right) \int_B P_u(dU)\right] \\
& \times \left[\prod_{i=1}^p \prod_{j=1}^{q_2} \exp\left[\frac{(\sigma_i b_i^j g_j(s) \int_B P_u(dU))^2}{2}\right]\right] \tag{2.16}
\end{aligned}$$

■

Las restricciones impuestas no son fuertes ni difíciles de cumplir ya que es habitual que las componentes principales tengan distribución Normal conjunta. Dado que son incorreladas, esto las hace independientes y por tanto normalmente distribuidas.

Como consecuencia de la proposición anterior, y bajo las mismas condiciones, se puede expresar la función de supervivencia condicionada como

$$\begin{aligned}
& P_{w/w_{n-1}}^B(W_n > s/W_{n-1}) \\
& \simeq \exp\left[\left(\Lambda^{q_1}(W_{n-1}) - \mu_\Lambda^2(s)\right) \int_B P_u(dU)\right] \\
& \times \exp\left[\frac{(\int_B P_u(dU))^2}{2} \sum_{j=1}^{q_2} \left(\sum_{i=1}^p (\sigma_i b_i^j)^2\right) (g_j(s))^2\right] \tag{2.17}
\end{aligned}$$

2.5.3. Contraste de bondad de ajuste

En el capítulo anterior, dedicado al PC, ya se desarrolló el contraste de bondad de ajuste de un PC en la Sección 1.4.3. En este apartado, se extiende ahora el estudio de la bondad de ajuste al caso del PC compuesto con marcas específicas.

Sea $\{N(t); t \geq t_0\}$ un PC compuesto. El estudio de sucesos de un proceso con marcas en un subconjunto dado B , es decir $\{N(t, B); t \geq t_0\}$, es de habitual interés. Utilizando los teoremas de representación, resulta que la intensidad de un PC compuesto con marcas específicas es nueva, pero sin embargo, el proceso de recuento continua siendo un PC, por lo que la generalización es clara.

Se tiene que el proceso $\{N(t, B); t \geq t_0\}$ es un PC con intensidad $\{\lambda^B(t) = \lambda(t) \int_B P_u(dU); t \geq t_0\}$. Por lo que tomando nuevas trayectorias de un nuevo proceso puntual $M(t)$, se puede comprobar si el proceso formado por puntos con marcas en B , $\{M(t, B); t \geq t_0\}$, es o no un PC con intensidad $\{\lambda^B(t) = \lambda(t) \int_B P_u(dU); t \geq t_0\}$.

De modo semejante al caso del PC, el contraste se establece de dos modos distintos, dependiendo de si hay conocimiento del proceso a contrastar en la hipótesis nula, o solo de los datos observados. Por tanto el test de bondad de ajuste sería

$$\begin{cases} H_0^{(j)} : E[\lambda^B(t_j)] = \mu(t_j) \int_B P_u(dU) \\ H_1^{(j)} : E[\lambda^B(t_j)] \neq \mu(t_j) \int_B P_u(dU) \end{cases}, \quad j = 1, \dots, p$$

en el caso en el que la intensidad del PC compuesto es conocida, independientemente de sus marcas. Obsérvese que $\sigma_{\lambda^B}^2(t) = \left[\int_B P_u(dU) \right]^2 \sigma^2(t)$. Siendo

el p-valor para el instante t_j

$$p - \text{valor}(t_j) = 2P \left[Z > \left| \frac{\mu_{exp}(t_j) - \int_B P_u(dU)\mu(t_j)}{\int_B P_u(dU)\sqrt{\frac{1}{s}\sigma^2(t_j)}} \right| \right]$$

En el caso en el que la estructura estocástica del proceso no es conocida y por tanto se realiza el test de bondad de ajuste a partir de los datos observados, se tiene el siguiente contraste

$$\begin{cases} H_0^{(j)} : E[\lambda^B(t_j)] = \hat{\mu}(t_j) \int_B P_u(dU) \\ H_1^{(j)} : E[\lambda^B(t_j)] \neq \hat{\mu}(t_j) \int_B P_u(dU) \end{cases}, \quad j = 1, \dots, p$$

en este caso la intensidad debe ser estimada entre los puntos con marcas en B , obteniéndose el estimador puntual de la intensidad (1.34) como ya se vió en la sección 1.4.3. De modo más explícito, se tiene

$$\hat{\lambda}_l^B(t_j) = \frac{1}{m} \sum_{i=1}^m \frac{M_i(t_j, B) - M_i(t_{j-1}, B)}{t_j - t_{j-1}}, \quad j = 1, \dots, p; \quad l = 1, \dots, s$$

Obsérvese que ahora $\hat{\sigma}_{\lambda^B}^2(t) = \left[\int_B P_u(dU) \right]^2 \hat{\sigma}^2(t)$ y el p-valor del contraste en t_j es

$$p - \text{valor}(t_j) = 2P \left[t > \left| \frac{\mu_{exp}(t_j) - \int_B P_u(dU)\hat{\mu}(t_j)}{\int_B P_u(dU)\sqrt{\frac{1}{s}\hat{\sigma}^2(t_j)}} \right| \right]$$

De este modo, queda extendido el test de bondad de ajuste al caso de un PC compuesto con marcas específicas en un subconjunto dado del espacio de marcas.

El criterio de rechazo de la hipótesis nula se hace siguiendo el criterio BH que se introdujo en el caso del PC (ver epígrafe 1.4.3).

2.6. Proceso de Cox multidimensional

Dejando de restringir el espacio de las ocurrencias del proceso puntual a la recta real, el PC se puede extender al caso en el que los sucesos pertenezcan a un espacio de varias dimensiones, teniendo así un *proceso de Cox multidimensional*. Ejemplos del estudio del PC multidimensional aplicado a la modelización de datos reales, pueden encontrarse en Ogata, 1998 [92], Spencer et al., 2000 [118], Haas y Shapiro, 2002 [62], Faenza et al., 2008 [49], Renshaw, 2009 [99], Myllymaki y Pettinen, 2010 [84], entre otros. Dentro de los PC multidimensionales pueden distinguirse dos casos diferenciados dependiendo de algunas características, estos son el *PC multichannel* y el *PC espacio-tiempo*.

Las aportaciones de esta tesis no parecen a priori aplicables a PC multidimensionales. Sin embargo, el hecho de que puedan interpretarse como PC compuestos, hace que muchos de ellos sean casos particulares de los tratados en los resultados desarrollados en esta memoria y por eso los mencionamos e interpretamos como compuestos en esta sección.

2.6.1. Procesos de Cox multichannel

Un PC multidimensional definido sobre un dominio espacio-temporal se denomina PC multichannel cuando el dominio espacial consiste en regiones discretas sobre las que se observa un PC temporal, mientras que las localizaciones espaciales dentro de las regiones son ignoradas. Desde un punto de vista matemático, podría interpretarse como un vector del PC, donde cada proceso asociado a una de las regiones se denomina *componente* del proceso multichannel. En Snyder y Miller, 1991 [117] pueden encontrarse ejemplos de la aplicación de los PC multichannel, como por el ejemplo el caso de un sis-

tema de comunicación óptica de detección directa, en el que las turbulencias atmosféricas hacen que el campo de luz recibido en el sistema sea espacialmente incoherente. Este efecto se ser compensado en algunos casos mediante el uso de una serie de fotodetectores en el receptor. Cada componente de un proceso multichannel puede asociarse con la salida de cada uno de los fotodetectores en la serie.

El modelo propuesto por Snyder y Miller, 1991 [117] para un PC multichannel, consta de K PC unidimensionales, donde cada uno sería una componente, de modo que son procesos mutuamente independientes dado el proceso información.

Definición 2.6.1 *Sea $\{N_k(t) : t \geq t_0\}$, $k = 1, \dots, K$ el PC con intensidad $\{\lambda_k(t, x(t)) : t \geq t_0\}$ donde $\{x(t) : t \geq t_0\}$ es el proceso información. Se asume que $N_1(\cdot), N_2(\cdot), \dots, N_K(\cdot)$ son mutuamente independientes dado el proceso información. El vector $\mathbf{N}(\cdot)$ con componentes $N_1(\cdot), N_2(\cdot), \dots, N_K(\cdot)$ se denomina PC multichannel con proceso intensidad*

$$\lambda(t, x(t)) = [\lambda_1(t, x(t)), \lambda_2(t, x(t)), \dots, \lambda_K(t, x(t))]'$$

Para las pretensiones de esta memoria, es importante destacar que este tipo de procesos pueden considerarse como un caso particular del PC compuesto en el que la marca de punto o suceso indique la región en la que éste ocurre.

2.6.2. Procesos de Cox espacio-tiempo

Existen diversos casos de PC multidimensionales en los que los puntos ocurren aleatoriamente en el espacio-tiempo, es decir, suceden a lo largo del tiempo y en una región del plano o el espacio con sus correspondientes

coordenadas. Por ejemplo, en medicina nuclear donde hay cámaras de rayos gamma, se utiliza para monitizar la radiactividad, cada fotón detectado tiene una coordenada espacial y otra temporal correspondiente al instante y localización de su absorción en el detector. Los seísmos ocurren aleatoriamente en el espacio y el tiempo sobre la superficie terrestre, provocando descargas luminosas, tormentas y otros fenómenos.

En Snyder y Miller, 1991 [117] presentan el siguiente modelo.

Definición 2.6.2 *Sea $N(T \times A)$ el número de puntos en $T \times A$, donde $T \times A$ es un conjunto medible en el dominio espacio-tiempo $T \times \mathcal{U}$ del proceso. Donde $T = [t_0, t[$ y para abreviar se denota $N(T \times A) = N(t, A)$. La intensidad del proceso se denota por $\{\lambda(t, \mathbf{u}, \mathbf{x}(t)), t \in \mathcal{T}, \mathbf{u} \in \mathcal{U}\}$ con $\{\mathbf{x}(t) : t \geq t_0\}$ el proceso información. Por definición $\mathbf{N}(t)$ es un PC espacio-tiempo sobre $\mathcal{T} \times \mathcal{U}$ si, condicionado al proceso información, es un proceso de Poisson espacio-tiempo.*

Trasladándolo al caso del PC compuesto, se observa que un proceso puntual en el cual los puntos tienen las coordenadas espacial y temporal, puede verse como un proceso marcado en el tiempo, donde las marcas de un punto son continuas e indican la posición espacial del mismo. Obteniéndose así que este proceso es un caso particular del PC compuesto.

Un PC espacio-tiempo también puede ser interpretado tanto como el límite de un PC multichannel, como el número de componentes del proceso incrementado hacia el infinito.

Capítulo 3

Proceso auto-excitado

Los capítulos anteriores de esta memoria se han dedicado a procesos que pretenden modelizar fenómenos puntuales de la forma más fiel posible aportando flexibilidad con respecto a su intensidad, incluso relajando la propiedad de regularidad con los procesos compuestos, pero sin considerar que el proceso tenga memoria. Es evidente que resultaría importante llevar esto a un nivel más elevado, permitiendo que la intensidad del proceso puntual pueda depender del pasado del propio proceso, como es el caso del proceso del que se ocupa el presente capítulo, el *proceso auto-excitado*. En este proceso, se pierde la propiedad de que el proceso no tenga memoria del PC pero se gana en generalización del modelo.

El estudio de procesos estocásticos con memoria no es ni mucho menos nuevo, piénsese por ejemplo en colas con memoria o series temporales. La novedad de los procesos auto-excitados es la posible dependencia de todo el pasado del proceso así como su doble aleatoriedad, es decir, que su intensidad pueda ser a su vez un proceso estocástico (Snyder y Miller, 1991 [117]). Kopperschmidt y Stute, 2013 [75] explica de forma muy clara y breve que los procesos de Hawkes fueron los primeros en aunar estas dos características;

además señala que es habitual la asunción de una forma multiplicativa para la intensidad por sus buenas propiedades asintóticas.

En casos en los que la intensidad sea más general, hay que subrayar la complejidad de sus expresiones y, por tanto, de su estimación. De hecho, los ejemplos de este tipo de procesos que se pueden encontrar en la literatura no son finalmente tan generales como se podría esperar.

Centrándonos en el proceso en sí, se puede observar que en el proceso auto-excitado más general, todo el pasado del proceso, es decir el número de ocurrencias $N(t)$ y los tiempos de ocurrencias $\omega_1, \omega_2, \dots, \omega_{N(t)}$ para todo punto anterior al tiempo t , puede influir en el número y tiempo de ocurrencia de todos los puntos posteriores. Para verlo de un modo más práctico, digamos que un ejemplo de fenómeno que puede ser modelado mediante un proceso auto-excitado es el de los isótopos radiactivos de vida corta, la intensidad de la radiación no será la misma al inicio de la vida del isótopo que transcurrido un cierto tiempo. Otro fenómeno que se podría modelar mediante el proceso auto-excitado es la emisión de electrones en un tubo de vacío en condiciones limitadas. En este caso, las emisiones desde el cátodo son suprimidas mediante los electrones emitidos previamente y que aún se encuentran en tránsito hacia el ánodo. De hecho, esta nube de electrones puede tener una carga total de una magnitud suficiente tal que la emisión de electrones puede estar casi extinguida. Esta emisión de electrones se modela mediante un proceso auto-excitado, como se puede ver de modo más detallado en Davenport y Root, 1958 [43] y en Srinivasan y Vasudevan, 1971 [120].

En la literatura, son muchos los artículos que tienen como base este tipo de procesos. Citando algunos, Hawkes, 1971 [65] hace un estudio teórico de los mismos en el que la intensidad está determinada por los sucesos en el pasado, haciendo distinción entre los efectos que puede tener: un efecto local

donde la función decrecería rápidamente o bien con efectos más a largo plazo que generaría una gráfica a partir del origen y con forma de joroba. Spreij, 1990 [119] se centra en un sistema de procesos de recuento auto-excitados con un espacio de estado finito, Al Mutairi, 1998 [5] utiliza los procesos auto-excitados para el estudio de la fiabilidad de software informático al igual que Wang, 2005 [124], y Hautsch, 2012 [69] que dedica una sección de su libro al caso particular de procesos auto-excitados estudiados por Hawkes y que toman su nombre, otros ejemplos interesantes pueden encontrarse en Mohler, 2011 [85] y Lewis, 2012 [79].

En Snyder y Miller, 1991 [117], se puede encontrar un estudio detallado de los procesos auto-excitados, en él dedican un capítulo al desarrollo de los mismos, de su expresión a partir de los tiempos de ocurrencia, y de la expresión de la función de densidad muestral que es base del estudio de otra generalización de los procesos de Poisson, los procesos de Cox y que se han visto anteriormente en el Capítulo 1 de esta memoria.

Sin embargo, hay muchas aplicaciones, en las que el proceso puntual no depende de la totalidad de su pasado, sino sólo de alguna pequeña fracción como puede ser el número total de puntos acumulado o de algunos de los últimos tiempos de ocurrencia. Los procesos más conocidos de este tipo son dos procesos muy estudiados, los *procesos de nacimiento de Markov* y los *procesos de renovación*.

A lo largo de este capítulo se definirán los procesos auto-excitados, se mostrarán los estadísticos de recuento de dichos procesos como son la función masa de probabilidad y algunos de sus momentos, para posteriormente pasar a la expresión de los principales estadísticos de tiempo. La última sección de este capítulo se centra en el estudio de la inferencia sobre el proceso y el proceso intensidad. La aportación clave de esta memoria en el presente

capítulo, es la estimación de la intensidad condicionada al recuento, a partir de la cual se podrán estimar los estadísticos de recuento del proceso auto-excitado.

3.1. Definición

Previo a la introducción de la definición del proceso auto-excitado, se establecen notación, terminología, así como ciertos supuestos siguiendo las pautas establecidas en Snyder y Miller, 1991 [117].

Sea $\{N(t) : t \geq t_0\}$ un proceso de recuento, y se asume que

1. $\{N(t) : t \geq t_0\}$ es condicionadamente regular.
2. El límite de la siguiente función $a(\Delta t, N(t))$ existe cuando Δt tiende a cero para casi toda realización de $\{N(t); t \geq t_0\}$

$$a(\Delta t, N(t)) = \begin{cases} \frac{1}{\Delta t} P [N(t, t + \Delta t) = 1/N(t)]; & N(t) = 0 \\ \frac{1}{\Delta t} P [N(t, t + \Delta t) = 1/N(t; \omega_1, \omega_2, \dots, \omega_{N(t)})]; & N(t) \geq 1 \end{cases}$$

donde ω_i es el tiempo de ocurrencia del i -ésimo punto.

3. $P [N(t_0) = 0] = 1$.

La primera propiedad indica que la probabilidad condicionada de dos o más puntos en un intervalo $(t, t + \Delta t)$ puede hacer una pequeña fracción arbitraria, de que haya solo un punto si se elige un Δt suficientemente pequeño. Se

denotan los límites de la segunda propiedad del siguiente modo

$$\begin{aligned}\mu(t, 0) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} P [N(t, t + \Delta t) = 1/N(t)]; \quad N(t) = 0 \\ \mu(t, N(t); \omega_1, \omega_2, \dots, \omega_{N(t)}) &= \\ \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} P [N(t, t + \Delta t) = 1/N(t); \omega_1, \omega_2, \dots, \omega_{N(t)}] &; \quad N(t) \geq 1\end{aligned}$$

estos caracterizan el comportamiento infinitesimal del proceso puntual condicionado a su comportamiento global previo. Se define el proceso intensidad como

$$\lambda(t) = \begin{cases} \mu(t, 0); & t_0 \leq t \leq \omega_1 \\ \mu(t, N(t); \omega_1, \omega_2, \dots, \omega_{N(t)}); & \omega_{N(t)} \leq t \leq \omega_{N(t)+1} \end{cases} \quad (3.1)$$

la cual es continua a la izquierda y se asume que $E[\lambda(t)] < \infty$, $t \geq t_0$.

Dada la definición de del proceso intensidad $\{\lambda(t); t \geq t_0\}$, se observa que es un proceso estocástico y que cualquiera de sus trayectorias muestrales depende del pasado del proceso de recuento.

El proceso de recuento $\{N(t); t \geq t_0\}$ se denomina proceso auto-excitado si su proceso intensidad depende de más que del tiempo t , de lo contrario, se trata de un proceso de Poisson no homogéneo. Si el proceso intensidad depende sólo del número de puntos en el intervalo $[t, t_0)$ y no de los tiempos de ocurrencia, el proceso $\{N(t); t \geq t_0\}$ se trata de un proceso de nacimiento. De modo general, la intensidad puede depender de uno o más tiempos de ocurrencia, así como del número de puntos anteriores.

3.1.1. Intensidad condicionada al recuento

Como puede verse en Snyder y Miller, 1991 [117], se define la *intensidad condicionada al recuento* como la intensidad esperada dado el número de puntos. Definida de modo más formal

$$\lambda_{N(t)}(t) = E[\lambda(t) \mid N(t)] \quad (3.2)$$

es decir, la esperanza cuando $N(t) = n$.

Otro modo de escribir la intensidad condicionada al recuento es como la intensidad vista en la definición en (3.1) pero condicionando solo a los puntos $N(t)$, tal y como se puede ver en Pham, 2003 [96]. Es claro que es una función del tiempo para cada $N(t) = n$, $n \in \mathbb{N}$.

Si se presta atención a la expresión de la definición de la intensidad (3.1) y se condiciona solo a $N(t) = n$, claramente la función $\lambda_{N(t^-)=n}(t)$ es distinta de cero en el intervalo donde es posible que antes de t haya n puntos. Además, $\lambda_{N(t)=n}(t)$ es distinta de cero entre el primero y el último de los posibles $n+1$ tiempos de ocurrencias.

Con el fin de simplificar la notación se usará en adelante $\lambda_n(t)$ en lugar de $\lambda_{N(t)=n}(t)$ para un n fijo.

Dada la definición del proceso intensidad condicionada, es fácil ver que

$$E_N[\lambda_{N(t)}(t)] = E[\lambda(t)] \quad (3.3)$$

que es la media de la intensidad del proceso auto-excitado expresada en función de la intensidad de recuento condicionada.

La importancia del proceso intensidad de recuento condicionada reside en que es un proceso paramétrico del proceso auto-excitado. Por lo tanto, no solo la intensidad está relacionada con él, sino que la función masa de probabilidad depende también de ella, así como otros momentos. Es muy

difícil, por no decir imposible, concretar cómo la intensidad depende tanto de los puntos $N(t)$ como de los tiempos de ocurrencias $\omega_1, \omega_2, \dots, \omega_{N(t)}$. Esto también sucede a la hora de dar una expresión cerrada de la intensidad condicionada al en un ejemplo real, incluso aunque solo dependa de $N(t)$, por lo que su estimación es una tarea interesante.

3.2. Estadísticos de recuento

Análogamente a los capítulos anteriores, esta sección se dedica al desarrollo de los estadísticos de recuento del proceso auto-excitado. La caracterización de dichos estadísticos se lleva a cabo mediante el estudio de la distribución del número de puntos en el intervalo $[t, t_0)$. Como se acaba de apuntar, el proceso intensidad condicionada al recuento definido anteriormente, tiene un papel muy relevante en el desarrollo de estos estadísticos.

3.2.1. Función masa de probabilidad

En primer lugar, se tiene que la función masa de probabilidad de un proceso auto-excitado tiene la siguiente expresión

$$\begin{aligned}
 P[N(t) = 0] &= \exp \left[- \int_{t_0}^t \lambda_0(\sigma) d\sigma \right] \\
 P[N(t) = n] &= \int_{t_0}^t \lambda_{n-1}(\tau) P[N(\tau) = n - 1] \exp \left[- \int_{\tau}^t \lambda_n(\sigma) d\sigma \right] d\tau
 \end{aligned} \tag{3.4}$$

según se demuestra en Snyder y Miller, 1991 [117].

La función masa de probabilidad viene dada mediante una ecuación recursiva que depende en cada paso de la intensidad condicionada. Esta expresión no es práctica ni para su aplicación a un caso real ni para su implementación. Sin embargo, el estimador de la misma que se propondrá en esta memoria

permitirá escribir una estimación de la función masa de probabilidad práctica teniendo en cuenta la información disponible en el caso de datos de sucesos recurrentes.

3.2.2. Momentos

Debido a la complejidad de la obtención de expresiones explícitas o incluso la estimación de los momentos del proceso auto-excitado, estos no son muy estudiados en la literatura. Por ello, y siguiendo las sugerencias de Snyder y Miller, se proporcionan expresiones para los principales momentos de estos procesos de recuento. Establezcamos previamente una proposición

Proposición 3.2.1 *Sea $\{N(t); t \geq t_0\}$ un proceso auto-excitado con proceso intensidad definido en (3.1) el cual es condicionalmente regular y además $E[\lambda(t)] < \infty$ y $P[N(t_0) = 0] = 1$, entonces el momento r -ésimo del número de puntos del proceso verifica*

$$\frac{\partial E[N^r(t)]}{\partial t} = \sum_{j=1}^r \binom{r}{j} E_N [N^{r-j}(t) \lambda_{N(t)}(t)], \quad r \in \mathbb{N} \quad (3.5)$$

Demostración. Bajo estos supuestos, se sabe que la probabilidad hasta un tiempo t , tomando n puntos, verifica

$$\begin{aligned} \frac{\partial P[N(t) = 0]}{\partial t} &= -\lambda_0(t)P[N(t) = 0]; \quad n = 0 \\ \frac{\partial P[N(t) = n]}{\partial t} &= -\lambda_n(t)P[N(t) = n] + \lambda_{n-1}(t)P[N(t) = n-1]; \quad n \geq 1 \end{aligned}$$

Multiplicando por n^r y sumando n , en la parte izquierda de la ecuación se obtiene que

$$\frac{\partial \sum_{n=0}^{\infty} n^r P[N(t) = n]}{\partial t} = \frac{\partial E_N[N^r(t)]}{\partial t}$$

y en la parte derecha

$$\begin{aligned}
& - \sum_{n=0}^{\infty} n^r \lambda_n(t) P[N(t) = n] + \sum_{n=0}^{\infty} n^r \lambda_{n-1}(t) P[N(t) = n-1] \\
& = -E_N [N^r(t) \lambda_{N(t)}(t)] + \sum_{n=0}^{\infty} (n-1+1)^r \lambda_{n-1}(t) P[N(t) = n-1] \\
& = -E_N [N^r(t) \lambda_{N(t)}(t)] + \sum_{n=0}^{\infty} \sum_{j=0}^r \binom{r}{j} (n-1)^r \lambda_{n-1}(t) P[N(t) = n-1] \\
& = -E_N [N^r(t) \lambda_{N(t)}(t)] + \sum_{j=0}^r \binom{r}{j} E_N [N^{r-j}(t) \lambda_{N(t)}(t)] \\
& = \sum_{j=1}^r \binom{r}{j} E_N [N^{r-j}(t) \lambda_{N(t)}(t)]
\end{aligned}$$

Por tanto,

$$\frac{\partial E[N^r(t)]}{\partial t} = \sum_{j=1}^r \binom{r}{j} E_N [N^{r-j}(t) \lambda_{N(t)}(t)]$$

como queríamos demostrar. ■

Caso $r = 1$: Media.

Tomando la ecuación (3.5) anterior para el caso $r = 1$:

$$\frac{\partial E[N(t)]}{\partial t} = E_N [\lambda_{N(t)}(t)] = \sum_{n=0}^{\infty} \lambda_n(t) P[N(t) = n] \quad (3.6)$$

e integrando entre t_0 y t ,

$$\int_{t_0}^t \frac{\partial E[N(\tau)]}{\partial \tau} d\tau = E[N(t)] = \sum_{n=0}^{\infty} \int_{t_0}^t \lambda_n(\tau) P[N(\tau) = n] d\tau \quad (3.7)$$

se obtiene la media del proceso auto-excitado en función de la intensidad condicionada al recuento.

Y como consecuencia de las ecuaciones (3.3), (3.6) y (3.7), vistas anteriormente, se obtiene también que

$$E[N(t)] = \int_{t_0}^t E[\lambda(\tau)] d\tau = \int_{t_0}^t E_N[\lambda_N(\tau)] d\tau$$

3.3. Estadísticos de tiempo

A lo largo de esta sección se obtienen varias expresiones asociadas a los estadísticos de los tiempos de ocurrencia de un proceso auto-excitado. El interés se centra en la *probabilidad de supervivencia*, la *densidad de ocurrencia conjunta* y la *función de densidad muestral*, siendo esta última la de mayor importancia.

3.3.1. Probabilidad de supervivencia

Sean $\omega_1, \omega_2, \dots$ los tiempos de ocurrencia de un proceso auto-excitado con proceso intensidad $\{\lambda(t); t \geq t_0\}$. Se denota por

$$\mathcal{P}_{\omega_{n+1}|\omega_1, \omega_2, \dots, \omega_n}(t | W_1, W_2, \dots, W_n)$$

la probabilidad condicionada de que se de una nueva ocurrencia en el instante ω_{n+1} , siendo $\omega_{n+1} \geq t_0$ dado que $\omega_i = W_i$ para $i = 1, 2, \dots, n$. En el caso que, por ejemplo, los tiempos de ocurrencias corresponden con algún comportamiento anómalo de una máquina, entonces \mathcal{P} es la probabilidad condicionada de que la máquina sobreviva hasta un tiempo t antes de que sucedan $(n + 1)$ anomalías, dadas las anomalías ocurridas en los tiempos W_1, W_2, \dots, W_n . Por ello \mathcal{P} se denomina *probabilidad de supervivencia condicionada* para los $(n + 1)$ primeros puntos. Debido a la identidad de los sucesos $\{N(t) = 0\}$ y $\{\omega_1 \geq t_0\}$, la probabilidad de supervivencia para el primer tiempo de ocurrencia, se determina fácilmente a partir de la ecuación de la función masa de probabilidad (3.4) siendo

$$\mathcal{P}_{\omega_1}(t) = P[\omega_1 \geq t] = \exp \left[- \int_{t_0}^t \lambda_0(\sigma) d\sigma \right] \quad (3.8)$$

De un modo más general, la probabilidad de supervivencia condicionada para los $(n + 1)$ primeros puntos para $t \geq W_n$ y $n \geq 1$ viene dada por

$$\mathcal{P}_{\omega_{n+1}|\omega_1,\omega_2,\dots,\omega_n}(t | W_1, W_2, \dots, W_n) = \exp \left[- \int_{W_n}^t \lambda(\sigma) d\sigma \right] \quad (3.9)$$

Mediante unas sencillas operaciones como se puede ver en Snyder y Miller, 1991 [117] se obtiene la siguiente ecuación

$$\lambda_n(t) = \frac{P_{\omega_{n+1}|\omega_1,\dots,\omega_n}(t | W_1, \dots, W_n)}{\mathcal{P}_{\omega_{n+1}|\omega_1,\omega_2,\dots,\omega_n}(t | W_1, W_2, \dots, W_n)} \quad (3.10)$$

donde el numerador es la intensidad condicionada para los $(n + 1)$ primeros tiempos de ocurrencia. A partir de esto se puede deducir la relación existente entre la intensidad y la probabilidad de supervivencia:

$$\lambda_n(t) = - \frac{\partial \ln \mathcal{P}_{\omega_{n+1}|\omega_1,\omega_2,\dots,\omega_n}(t | W_1, W_2, \dots, W_n)}{\partial t} \quad (3.11)$$

a partir de esta ecuación se obtiene la expresión general de la probabilidad de supervivencia (3.9).

3.3.2. Función de densidad conjunta

Denotemos como $p_w^{(n)}(W)$ a la *función de densidad conjunta* para los n primeros tiempos de ocurrencias, la cual viene dada por

$$p_{\omega_1}^{(1)}(W_1) = \lambda_0(W_1) \exp \left[- \int_{t_0}^{W_1} \lambda_0(\sigma) d\sigma \right] \quad (3.12)$$

para $n = 1$ y

$$p_w^{(n)}(W) = \lambda_0(W_1) \left\{ \prod_{i=2}^n \lambda_{i-1}(W_i) \right\} \times \exp \left[- \int_{t_0}^{W_1} \lambda_0(\sigma) d\sigma - \sum_{i=2}^n \int_{W_{i-1}}^{\omega_i} \lambda_{i-1}(\sigma) d\sigma \right] \quad (3.13)$$

para $n \geq 2$ y $t_0 \leq W_1 < \dots = W_n$. La ecuación primera ecuación (3.12) se obtiene derivando la ecuación (3.8) de la probabilidad de supervivencia, de la subsección anterior, respecto de t , mientras que para poder demostrar la segunda expresión (3.13) es necesario utilizar la siguiente relación

$$p_w^{(n)} = p_{\omega_1}(W_1) \prod_{i=2}^n p_{\omega_i|\omega_1, \dots, \omega_{i-1}}(W_i | W_1, \dots, W_{i-1})$$

Las densidades de la derecha se obtienen derivando las ecuaciones de la probabilidad de supervivencia (3.8) y (3.9) respecto de t , llegando así a la ecuación (3.13).

3.3.3. Función de densidad muestral

La función de densidad muestral de un proceso puntual auto-excitado tiene un papel relevante tanto en la decisión como en la estimación de problemas en los que toda realización del proceso es observada. Se define la función de densidad muestral de un proceso auto-excitado como la probabilidad conjunta de un número de puntos en el intervalo $[t_0, t)$ y las densidades de sus tiempos de ocurrencia. Se denota por

$$p[\{N(\sigma); t_0 \leq \sigma < T\}] = \begin{cases} P[N(T) = 0], & N(T) = 0 \\ p_w[W, N(T) = n] & N(T) \geq 1 \end{cases} \quad (3.14)$$

donde

$$\begin{aligned} p_w[W, N(T) = n] &= P[N(T) = n | \omega_1 = W_1, \dots, \omega_n = W_n] p_w^{(n)} \\ &= \mathcal{P}_{\omega_{n+1}|\omega_1, \dots, \omega_n}(T | W_1, \dots, W_n) \end{aligned}$$

para $n \geq 1$.

Haciendo uso de las expresiones de la probabilidad de supervivencia (3.8) y de la función de densidad conjunta, (3.12) y (3.13) en la expresión de la

función de densidad muestral (3.14), se obtiene, por un lado,

$$p[\{N(\sigma); t_0 \leq \sigma < T\}] = \begin{cases} \exp \left[\int_{t_0}^T \lambda_0(\sigma) d\sigma \right], & N(T) = 0 \\ \lambda_0(W_1) \exp \left[\int_{t_0}^{W_1} \lambda_0(\sigma) d\sigma - \int_{W_1}^T \lambda_1(\sigma) d\sigma \right], & N(T) = 1 \end{cases} \quad (3.15)$$

para $N(T) = 0$ y $N(T) = 1$;

y por otro lado,

$$p[\{N(\sigma); t_0 \leq \sigma < T\}] = \lambda_0(W_1) \left\{ \prod_{i=2}^n \lambda_{i-1}(W_i) \right\} \times \exp \left[\int_{t_0}^{W_1} \lambda_0(\sigma) d\sigma - \sum_{i=2}^n \int_{W_{i-1}}^{W_i} \lambda_{i-1}(\sigma) d\sigma - \int_{W_n}^T \lambda_n(\sigma) d\sigma \right] \quad (3.16)$$

para $n \geq 2$ y $t_0 \leq W_1 < W_2 < \dots < W_n < T$.

Por tanto, la función de densidad conjunta se puede expresar en términos del proceso de recuento $\{N(t); t \geq t_0\}$ y de su proceso intensidad $\{\lambda(t); t \geq t_0\}$ como se resume en el siguiente teorema.

Teorema 3.3.1 *Función de densidad conjunta para un proceso puntual auto-excitado*

Sea $\{N(t); t \geq t_0\}$ un proceso de recuento auto-excitado con proceso intensidad $\{\lambda(t); t \geq t_0\}$. La función de densidad conjunta para $\{N(\sigma); t_0 \leq \sigma < T\}$ es

$$p[\{N(\sigma); t_0 \leq \sigma < T\}] = \exp \left[- \int_{t_0}^T \lambda(\sigma) d\sigma + \int_{t_0}^T \ln \lambda(\sigma) dN(\sigma) \right] \quad (3.17)$$

donde

$$\int_{t_0}^T \ln \lambda(\sigma) dN(\sigma) = \begin{cases} 0, & N(T) = 0 \\ \ln \lambda_0(\omega_1), & N(T) = 1 \\ \ln \lambda_0(\omega_1) + \sum_{i=2}^{N(T)} \ln \lambda_{i-1}(\omega_i), & N(T) \geq 2 \end{cases} \quad (3.18)$$

En Snyder y Miller, 1991 [117, pág 302] se pueden encontrar ejemplos acerca de la función de densidad muestral.

3.4. Estimador de la intensidad condicionada al recuento

Pasemos ahora a abordar el tema de la inferencia sobre los procesos auto-excitados. Para ello, tomemos varias trayectorias muestrales del proceso de recuento $N(t)$, $k = 1, \dots, K$, observadas en el intervalo $[t_0, T]$. En otras palabras, se han observado K sujetos a lo largo del tiempo hasta un instante T , usando la expresión habitual en otros ámbitos como en análisis de sucesos recurrentes, Aalen et al., 2008 [1]. Siguiendo con la misma línea de trabajo que se ha propuesto anteriormente en esta memoria, se partirá del desconocimiento de estructura o dependencia del pasado para la intensidad o para la intensidad de recuento condicionada, por lo que los datos observados son

la única información disponible. Por tanto, se tratará con los datos de los sucesos recurrentes registrados en las trayectorias muestrales consideradas y sus tiempos de ocurrencias. La notación usada será la siguiente. Se denota

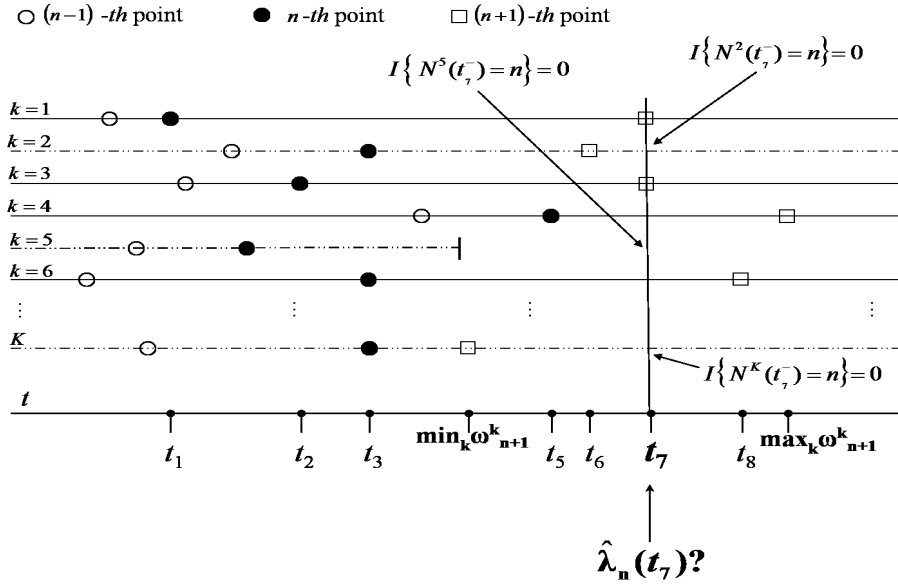


Figura 3.1: Diagrama de las trayectorias muestrales del proceso y sus puntos estimada la intensidad de recuento condicionada en un instante de tiempo.

por ω_i^k al tiempo de ocurrencia del i -ésimo punto de la k -ésima trayectoria muestral con $1 \leq i \leq p_k \leq T$, $k = 1, 2, \dots, K$ y $\omega_0^k = t_0$. De modo que las trayectorias pueden tener diferente longitud, surgiendo así la posibilidad de retirar un sujeto del estudio, o que este se dé de baja. Se denota por $N^k(t_j)$ o $N^k(\omega_{n-1}^k)$, al número de puntos hasta t_j o ω_{n-1}^k para la k -ésima trayectoria. En este caso, es importante hacer que la intensidad de recuento condicionada sea distinta de cero en $[\min_k \omega_{n+1}^k, \max_k \omega_{n+1}^k]$, para cada $n \in \mathbb{N}$. En este intervalo, se propone el siguiente estimador puntual:

$$\hat{\lambda}_n(t_j) = \frac{\sum_{k=1}^K \frac{N^k(t_j) - N^k(\omega_n^k)}{t_j - \omega_n^k} I \{N^k(t_j^-) = n\}}{\sum_{k=1}^K I \{N^k(t_j^-) = n\}}, \quad n \in \mathbb{N} \quad (3.19)$$

para $t_j < T$ y donde t_j^- es el tiempo al que tiende t_j por la izquierda, e I es la función indicadora .

Nótese que la intensidad condicionada al recuento, dado un número fijo de puntos, $\widehat{\lambda}_n(t_j) = \widehat{E}[\lambda(t_j) \mid N(t_j) = n]$, es una función del tiempo que ha sido estimada en un conjunto discreto de instantes de tiempo $t_j = \omega_{n+1}^k$. Dicho de otra manera, la ecuación (3.19) proporciona estimaciones puntuales en distintos instantes de tiempo t , para todas las funciones intensidad condicionadas al recuento ($\forall n \in \mathbb{N}$), en el intervalo de observación de las trayectorias muestrales.

En el caso de un proceso auto-excitado, las expresiones de sus estadísticos son especialmente complicadas de aplicar en casos reales por lo que cualquier particularización y el cálculo de una expresión práctica es siempre importante. El estimador propuesto en esta sección, servirá para estimar también algunos estadísticos de recuento del proceso auto-excitado.

3.4.1. Estimación de la función masa de probabilidad

La ecuación de la función masa de probabilidad (3.4), proporciona una expresión de ésta en función de la intensidad de recuento condicionada. Con el fin de aplicar el estimador propuesto en la ecuación (3.19), han de tenerse en cuenta dos detalles de gran importancia . En primer lugar, que $\widehat{\lambda}_n(t_j)$ es distinta de cero en el intervalo $[\text{mín}_k \omega_{n+1}^k, \text{máx}_k \omega_{n+1}^k]$, $n \in \mathbb{N}$ y, en segundo lugar, que los intervalos para cada $\widehat{\lambda}_n(t)$ y $\widehat{\lambda}_{n-1}(t)$ que sean distintos de cero, pueden superponerse.

Teniendo esto en cuenta, la función masa de probabilidad de la ecuación (3.4) puede estimarse como

$$\widehat{P}[N(t) = n] = \int_{\text{mín}_k \omega_n^k}^t \widehat{\lambda}_{n-1}(\tau) \widehat{P}[N(\tau) = n - 1] \exp \left[- \int_{\tau}^t \lambda_n(\sigma) d\sigma \right] d\tau$$

y depende del tiempo t , si son intervalos superpuestos

i. $\min_k \omega_n^k < t < \min_k \omega_{n+1}^k$:

$$\widehat{P}[N(t) = n] = \int_{\min_k \omega_n^k}^t \widehat{\lambda}_{n-1}(\tau) \widehat{P}[N(\tau) = n-1] d\tau$$

ii. $\min_k \omega_{n+1}^k \leq t < \max_k \omega_n^k$:

$$\begin{aligned} \widehat{P}[N(t) = n] = & \\ & \int_{\min_k \omega_n^k}^{\min_k \omega_{n+1}^k} \widehat{\lambda}_{n-1}(\tau) \widehat{P}[N(\tau) = n-1] d\tau \\ & + \int_{\min_k \omega_{n+1}^k}^t \widehat{\lambda}_{n-1}(\tau) \widehat{P}[N(\tau) = n-1] \exp \left[- \int_{\tau}^t \widehat{\lambda}_n(\sigma) d\sigma \right] d\tau \end{aligned}$$

y si los intervalos no se superponen, $\min_k \omega_n^k < t < \max_k \omega_n^k$, entonces

$$\widehat{P}[N(t) = n] = \int_{\min_k \omega_n^k}^t \widehat{\lambda}_{n-1}(\tau) \widehat{P}[N(\tau) = n-1] d\tau$$

Aplicando la estimación de cada $\widehat{\lambda}_n$ según la ecuación (3.19), se calcularía la estimación final de la función masa de probabilidad.

3.4.2. Estimación de la media

Siguiendo en el mismo escenario, k trayectorias muestrales en las que se han observado los tiempos de ocurrencia, la expresión de la media se puede extender con el fin de encontrar una forma cerrada de la misma. Teniendo en cuenta la ecuación (3.7), y que las funciones $\widehat{\lambda}_n(t)$ son distintas de cero en el intervalo $[\min_k \omega_{n+1}^k, \max_k \omega_{n+1}^k]$, $n \in \mathbb{N}$, la media puede estimarse como

$$\widehat{E}[N(t)] = \sum_{n=0}^{\infty} \int_{\min_k \omega_{n+1}^k}^t \widehat{\lambda}_n(\tau) \widehat{P}[N(\tau) = n] d\tau \quad (3.20)$$

El inconveniente que tiene esta expresión es que, siendo una suma de términos infinitos, no es útil para su cálculo en datos finitos. Sin embargo,

para un t fijo, se denota por ω_S al máximo extremo inferior de los intervalos en los que las funciones $\widehat{\lambda}_n(t)$ son distintas de cero. En este caso, S se define como

$$S = \underset{n}{\text{máx}}\{n : t \geq \underset{k}{\text{mín}}\omega_n^k\}$$

y verifica que

$$\widehat{\lambda}_{S+i}(\tau) = 0, \quad \tau \in [\underset{k}{\text{mín}}\omega_{n+1}^k, t] \quad \text{para } i = 0, 1, 2, \dots$$

A parte, la suma de la ecuación del estimador de la media anterior (3.20), tiene muchos términos que son cero, por lo que dicho estimador puede reescribirse como

$$\widehat{E}[N(t)] = \sum_{n=0}^{S-1} \int_{\underset{k}{\text{mín}}\omega_{n+1}^k}^t \widehat{\lambda}_n(\tau) \widehat{P}[N(\tau) = n] d\tau \quad (3.21)$$

Es conveniente observar que hay intervalos completos donde las funciones $\widehat{\lambda}_n(t)$ son distintas de cero y se encuentran antes del instante t . Se denota por ω_J al extremo superior máximo de dichos intervalos por lo que es menor que t y por tanto J se define como

$$J = \underset{n}{\text{máx}}\{n : t \geq \underset{k}{\text{máx}}\omega_n^k\}$$

Entonces también se verifica que

$$\widehat{\lambda}_i(\tau) \neq 0, \quad \tau \in [\underset{k}{\text{mín}}\omega_{n+1}^k, \underset{k}{\text{máx}}\omega_{n+1}^k] \quad \text{para } i = 0, \dots, J$$

y obviamente $J < S$.

Por lo que finalmente, se obtiene que la media de un proceso auto-excitado visto como datos de sucesos recurrentes puede expresarse como la suma de

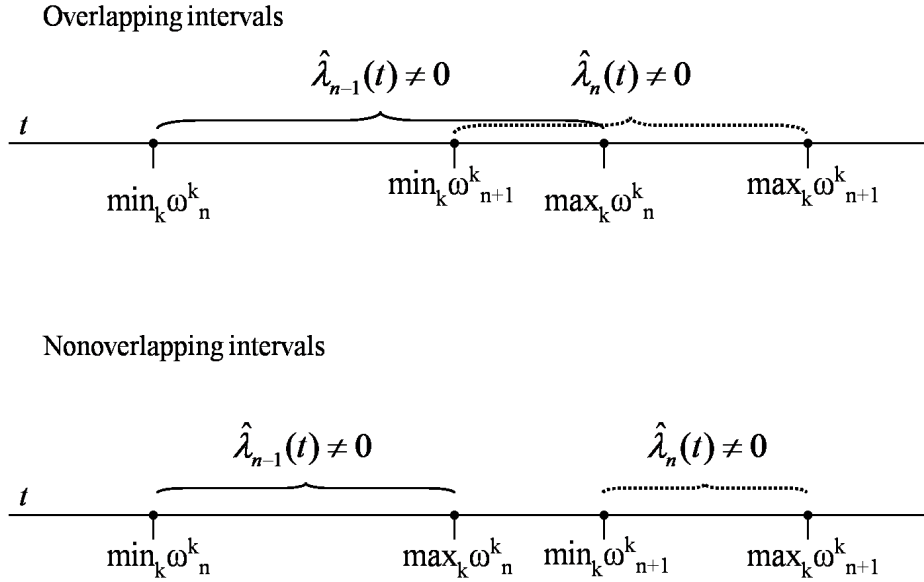


Figura 3.2: Diagrama del estimador de la intensidad de recuento condicionada dadas $n-1$ y n ocurrencias cuando los intervalos donde las funciones son distintas de cero se superponen o no.

un número finito de términos

$$\begin{aligned} \widehat{E}[N(t)] = & \\ & \sum_{n=0}^J \int_{\min_k \omega_{n+1}^k}^{\max_k \omega_{n+1}^k} \widehat{\lambda}_n(\tau) \widehat{P}[N(\tau) = n] d\tau + \\ & \sum_{n=J+1}^{S-1} \int_{\min_k \omega_{n+1}^k}^t \widehat{\lambda}_n(\tau) \widehat{P}[N(\tau) = n] d\tau \end{aligned} \quad (3.22)$$

3.4.3. Estimación de la intensidad media

La media del proceso intensidad viene dada en términos de la intensidad de recuento condicionada de la ecuación (3.3). Como consecuencia, el estimador de la intensidad de recuento condicionada y la función masa de probabilidad pueden utilizarse para estimar la media del proceso intensidad

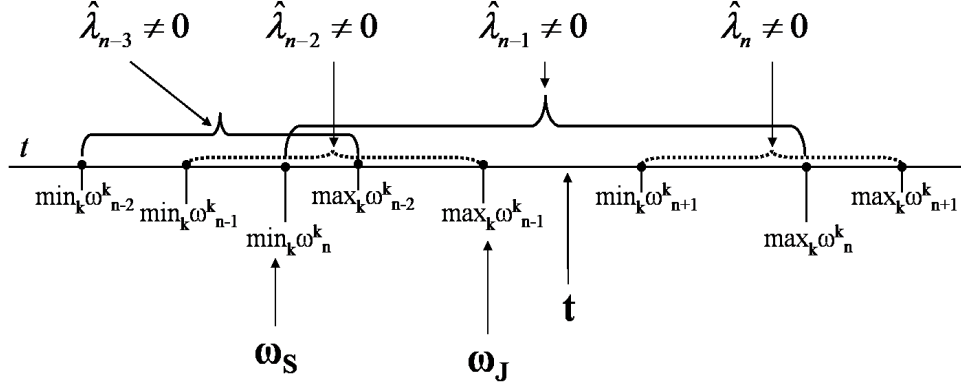


Figura 3.3: Diagrama para ilustrar los tiempos ω_S y ω_J .

del siguiente modo

$$\widehat{E}[\lambda(t)] = \widehat{E}_N[\lambda_{N(t)}(t)] = \sum_{n=0}^{\infty} \widehat{\lambda}_n(t) \widehat{P}[N(t) = n] \quad (3.23)$$

Utilizando una vez más donde las funciones $\widehat{\lambda}_n(t)$ son distintas de cero, la ecuación (3.23) queda de la siguiente manera

$$\widehat{E}[\lambda(t)] = \sum_{n=J+1}^S \widehat{\lambda}_n(t) \widehat{P}[N(t) = n] \quad (3.24)$$

Por último, recordar que el proceso auto-excitado es una generalización de proceso de Cox, en el cual existe una dependencia parcial o total del tiempo como ya se ha indicado en diversas ocasiones a lo largo del capítulo. Con el fin de que este hecho quede remarcado y se vea de un modo más gráfico, se incluye a continuación un teorema en el cual se caracteriza el PC como caso particular del proceso auto-excitado.

Teorema 3.4.1 *Caracterización de un PC como un proceso puntual autoexcitado*

Sea $\{N(t) : t \geq t_0\}$ un PC con proceso intensidad $\{\lambda(t, x(t))\}$, y se supone que $\lambda(\cdot)$ es integrable. Entonces $\{N(t) : t \geq t_0\}$ es un proceso auto-excitado

con proceso intensidad $\{\hat{\lambda}(t) : t \geq t_0\}$, donde

$$\begin{aligned} \hat{\lambda}(t) &= E[\lambda(t, x(t)) \mid N(\sigma) : t_0 \leq \sigma < t] \\ &= \begin{cases} E[\lambda(t, x(t)) \mid N(t)], & N(t) = 0 \\ E[\lambda(t, x(t)) \mid N(t); \omega_1, \dots, \omega_{N(t)}], & N(t) \geq 1 \end{cases} \end{aligned} \quad (3.25)$$

Utilizando la ecuación de la probabilidad de supervivencia de un proceso auto-excitado y en combinación con (3.25) se tiene que la probabilidad de supervivencia condicionada para los $(n + 1)$ primeros puntos de un PC para $t \geq W_n$ y $n \geq 1$ viene dada por

$$P_{\omega_{n+1}|\omega_1, \dots, \omega_n}(t \mid W_1, \dots, W_n) = \exp\left(-\int_{W_n}^t \hat{\lambda}(\sigma) d\sigma\right) \quad (3.26)$$

donde $\hat{\lambda}(\sigma) = E[\lambda(\sigma, x(\sigma)) \mid N(\sigma) = n; W_1, \dots, W_n]$. Utilizando la caracterización del teorema (3.4.1) se pueden obtener expresiones para la función de densidad de conjunta y para la función de densidad muestral como puede verse en Snyder y Miller, 1991 [117, pág. 365-366].

Capítulo 4

Simulaciones y Aplicaciones

Con el fin de ilustrar los resultados teóricos propuestos a lo largo de la memoria, en este capítulo se utilizarán tanto datos simulados para su tratamiento con los métodos desarrollados, como aplicaciones de éstos a datos reales procesados. El motivo por el cual se han unido en un capítulo tanto las simulaciones como las aplicaciones a datos reales, en lugar de situarlas a lo largo del documento, es porque en cada una de ellas se han utilizado resultados obtenidos en distintos capítulos. De esta manera, la presentación de los resultados resulta más completa y a su vez más clara al no estar subdividida en las distintas secciones anteriores, y se aprecian mejor todos los resultados que se pueden aplicar a un ejemplo. El software utilizado para construir las simulaciones, para la organización de los datos, así como para la aplicación de los resultados teóricos, ha sido MATLAB R2016b.

4.1. Simulaciones. Metodología.

En primer lugar, antes de abordar las simulaciones en sí, comenzamos describiendo el método de simulación que ha sido empleado en la realización

de las mismas, para poder continuar mostrando los distintos ejemplos. Es necesario tener en cuenta previamente, que de la gran variedad de simulaciones que se podrían obtener, por la elección de la distribución de la intensidad y de las marcas, se han elegido ejemplos distintos con el fin de ilustrar, por una parte, las metodologías de inferencia referidas a estadísticos de tiempo y de recuento y, por otra, la referida a bondad de ajuste.

Para simular las ocurrencias del PC compuesto, se ha comenzado simulando los tiempos de ocurrencia del PC, para posteriormente, a cada tiempo de ocurrencia, asignarle la correspondiente marca, dependiendo de la situación que se desee simular.

El procedimiento utilizado para la simulación de los tiempos de ocurrencia del PC ha sido el de la transformada inversa, teniendo en cuenta que los tiempos entre llegadas de un proceso de Poisson siguen una distribución exponencial. Este método es aplicable si se puede invertir de forma analítica la función de distribución de la variable aleatoria para la que deseamos generar nuevos valores, en este caso la función de distribución de una variable aleatoria exponencial.

Por tanto, si suponemos que $X = \omega_t - \omega_{t-1} \rightsquigarrow Exp(\lambda)$, para dos tiempos de ocurrencia sucesivos ω_{t-1} , ω_t , y tenemos en cuenta que su función de distribución es $u = F_X(x) = 1 - exp^{-\lambda x}$, $x \geq 0$, $\lambda > 0$, es necesario utilizar un generador de números aleatorios para obtener valores u de una variable aleatoria uniforme $U \rightsquigarrow U([0, 1])$ y con ello

$$u = 1 - exp^{-\lambda(\omega_t - \omega_{t-1})} \Rightarrow \omega_t = \omega_{t-1} - \frac{1}{\lambda} \log(1 - u)$$

Este procedimiento correspondería a la generación de tiempos de ocurrencia de un proceso de Poisson homogéneo de intensidad λ . Si ahora incorporamos la condición de aleatoriedad a la intensidad, se generan tiempos de ocurrencia de un PC y, posteriormente, se obtiene el recuento de puntos y

valores generados del proceso de recuento $\{N(t); t \geq 0\}$.

En esta memoria se utilizan distintas simulaciones del PC compuesto que corresponden a situaciones con intensidades aleatorias de distribución conocida, por ejemplo, intensidad Binomial, intensidad Gamma e intensidad log-normal. Sin embargo, también se simulan otras situaciones más complejas como es el caso de los procesos tipo fase en los cuales la intensidad no solo es una variable aleatoria tipo fase, sino que además su representación contiene parámetros aleatorios, o el caso de los procesos autoexcitados en los que se incorpora el pasado del proceso. En todos los casos el software proporciona solución a la ecuación que se deriva de la aplicación del método de la transformada inversa.

4.2. Simulaciones del PC compuesto

En esta sección se ofrecen tres simulaciones de PC compuestos con distintas intensidades y espacios de marcas. La selección de los mismos ha sido inspirada por ejemplos encontrados en distintos ámbitos de conocimiento. Los modelos más usuales tienen como intensidad una distribución gamma o un movimiento Browniano geométrico y como distribuciones para las marcas, distribuciones binomiales, gamma y lognormales (ver Gospodinov y Rotondi, 2001 [58], Barta et al. 2005 [10], Park y Padgett, 2005 [94] y, en general, artículos de la bibliografía). La distribución binomial es interesante por ejemplo, para modelizar el problema de posibles ocurrencias simultáneas o supresión de ocurrencias (ver Snyder y Miller, 1991 [117]). Sin embargo, los valores de los parámetros de las distribuciones no se han escogido por ninguna razón especial, son meros ejemplos y cualquier otro valor habría servido para el mismo propósito.

Las tres simulaciones se usarán para ilustrar el cálculo de los estadísticos de un PC compuesto. En cada simulación se considerará un número k de trayectorias muestrales en un intervalo de tiempo $[T_0, T_2]$. El intervalo se cortará en el instante T_1 para considerar el intervalo $(T_0, T_1]$ el intervalo del pasado y el intervalo (T_1, T_2) , el futuro, a la hora de la estimación del proceso media (ver ecuaciones (1.29) y (2.11)). A partir de dicha estimación, se estima también la moda (ver ecuación (2.12)) y la probabilidad de tener una nueva ocurrencia antes de un instante de tiempo $s \in (T_1, T_2)$. Para una mejor ilustración de los estadísticos citados, se han escogido distintos valores de T_1 y de s .

Estas estimaciones, no solo ilustran los resultados teóricos expuestos a lo largo de la memoria, sino que, en cada caso, se evalúa la precisión de los cálculos realizados.

El último caso de simulación, se utiliza también para ilustrar la aplicación del test de bondad de ajuste expuesto en el epígrafe 2.5.3 del Capítulo 2. Tanto los test unidimensionales, como su visión conjunta desde el punto de vista de Inferencia Simultánea y el criterio BH son conocidos por lo que no habría que evaluarlos. Sin embargo, es interesante observar el comportamiento del test de bondad de ajuste para el PC compuesto construido a partir del uso conjunto de todos ellos tal y como se ha propuesto en esta memoria. Más adelante se comentará su comportamiento en la aplicación a casos reales

4.2.1. Simulación PC caso 1

Esta simulación es un caso de PC compuesto que tiene como proceso intensidad una distribución Gamma(4,0.5), con marcas distribuidas según una Binomial(1,0.6) y como subconjunto de marcas se ha tomado $B = \{u; u = 1\}$ con probabilidad $\int_B P_u(dU) = p(1) = 0.6$. Es decir, este caso se puede inter-

pretar como un PC con supresiones aleatorias, los puntos con marca $u = 0$ son borrados, y solo se tienen en cuenta los puntos con marca $u = 1$.

Como se ha adelantado, se van a ilustrar los cálculos de la estimación de la media, la moda y la probabilidad de tener una nueva ocurrencia antes de un instante fijo del futuro. Para que resulte más claro, se harán estas estimaciones en tres escenarios distintos, concretamente en los intervalos futuros $(T_1 = 5, T_2 = 20]$, $(T_1 = 10, T_2 = 20]$ o $(T_1 = 15, T_2 = 20]$. Además de realizarlas en estos tres intervalos, también se han llevado a cabo para diversos instantes de tiempo s dentro de cada intervalo.

En esta simulación, el proceso media ha sido modelizado como un $PCP(q_2; p_1, \dots, p_j) \equiv PCP(6; 4, 4, 0, 4, 3, 4)$. A priori, el modelo implica que 6 componentes del futuro se utilizan en la estimación, y que cada una de ellas es explicada con 4, 4, 0, 4, 3 y 4 componentes del pasado respectivamente. No obstante, si se examinan los parámetros del modelo en detalle, es interesante notar que la tercera componente del futuro en realidad no está explicada por ninguna del pasado. Por tanto, solo 5 componentes del futuro se mantienen en el modelo. Por lo que habiendo computado los coeficientes y teniendo en cuenta tanto los coeficientes nulos como los no nulos, el proceso media para $s \in (T_1 = 5, T_2 = 20]$ puede escribirse del siguiente modo

$$\tilde{\Lambda}^6(s) = \mu_{\Lambda}^2(s) + \begin{pmatrix} -1.024 & 0.962 & -1.654 \\ -2.903 & 1.488 & -3.965 \\ -0.425 & 0.162 & -0.263 \\ -0.061 & 0.115 & 0 \\ -0.457 & 0.212 & -0.404 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_3 \\ \xi_4 \end{pmatrix} (g_1(s), g_2(s), g_4(s), g_5(s), g_6(s)).$$

En el caso de elegir otros intervalos del futuro distintos a los escogidos en este caso, deberían reiterarse estos mismos cálculos. Del mismo modo, cambiando

el instante de tiempo T_1 , se obtiene otro modelo PCP. Puntualizar que en el caso en que el instante T_1 sea muy cercano a T_0 , el intervalo del pasado sería muy pequeño, no teniendo suficiente información para poder estimar y predecir el futuro con buenos resultados. Por otro lado, para un T_1 fijo, cuanto más lejana sea la estimación de s , mayor será el error de estimación esperado. Por lo que puede decirse que el error de estimación tiende a incrementarse a medida que s se acerca al límite superior del intervalo futuro, T_2 .

En la tabla 4.1 se muestran los resultados obtenidos de las estimaciones de la media, la moda y de la probabilidad de tener una nueva ocurrencia antes de un instante fijo del futuro, correspondientes a este ejemplo. Como ya se indicó con anterioridad, se realizan en tres escenarios distintos y para diferentes instantes s dentro de cada intervalo.

Aunque no sea necesario probar la precisión de la estimación del proceso media dada en la ecuación (1.29), puesto que como se comentó en el Capítulo 1, en Bouzas et al., 2006c [23] ya se comprobó el pequeño error de estimación que se consigue con el método utilizado, sin embargo es ilustrativo comparar las estimaciones dadas en la tabla 4.1 con las medias muestrales del proceso $N(t, B)$ en el mismo instante s , calculado de todas las trayectorias muestrales simuladas y que se presentan en la tabla 4.2. Como puede observarse, la comparación entre ambas tablas muestra la similitud entre los resultados obtenidos en los dos casos o el pequeño error que los diferencia.

Con el fin de estudiar el error de estimación de la moda y de la probabilidad de tener una nueva ocurrencia antes de un tiempo s , se simularon 100 nuevas trayectorias muestrales. La precisión de estas estimaciones puede evaluarse para cada instante de tiempo T_1 que divide el intervalo de observación en el pasado y el futuro, y también para cada instante s del intervalo del futuro. Sin embargo, en esta sección se presenta el estudio de los errores

s	$T_1 = 5$			$T_1 = 10$			$T_1 = 15$		
	E[N(s)]	Moda	Prob	E[N(s)]	Moda	Prob	E[N(s)]	Moda	Prob
6	7.217	6	7	0.511					
7	8.438	7	8	0.865					
8	9.651	8	9	0.959					
9	10.865	9	10	0.988					
10	12.072	11	12	0.996					
11	13.300	12	13	0.998	13.300	12	13	0.661	
12	14.496	13	14	0.999	14.496	13	14	0.904	
13	15.685	14	15	0.999	15.685	14	15	0.971	
14	16.900	15	16	0.999	16.900	15	16	0.991	
15	18.109	17	18	0.999	18.109	17	18	0.997	
16	19.286	18	19	0.999	19.286	18	19	0.999	19.286
17	20.492	19	20	0.999	20.492	18	19	0.999	20.492
18	21.684	20	21	0.999	21.684	20	21	0.999	21.684
19	22.884	21	22	0.999	22.884	21	22	0.999	22.884
20	24.061	23	24	0.999	24.061	23	24	0.999	24.061

Tabla 4.1: Simulación PC 1. Estimación de la media, moda y probabilidad de tener una nueva ocurrencia antes de un tiempo s (Prob) del proceso $N(t, B)$.

s	Estimación de la media			Medias muestrales
	$T_1 = 5$	$T_1 = 10$	$T_1 = 15$	
6	7.217			7.224
7	8.448			8.450
8	9.651			9.675
9	10.865			10.890
10	12.072			12.102
11		13.300		13.292
12		14.496		14.512
13		15.685		15.727
14		16.900		16.918
15		18.109		18.123
16			19.286	19.320
17			20.492	20.520
18			21.684	21.733
19			22.884	22.919
20			24.061	24.117

Tabla 4.2: Simulación PC 1. Comparativa de las medias estimadas y las medias muestrales para cada trayectoria del proceso $N(t, B)$.

para el caso concreto en el que $T_1 = 10$ y $s = 11, 15$ y 20 puesto que sería demasiado amplio presentarlo para todo T_1 y todo s .

Los porcentajes muestrales del número de puntos antes de un instante de tiempo s se han calculado a partir de las 100 nuevas trayectorias muestrales. La tabla 4.3 muestra, para los tres valores de s citados, los valores más frecuentes así como sus frecuencias. Si se comparan los valores obtenidos en la tabla 4.3 con las estimaciones de la moda dadas en la tabla 4.1, se puede observar que la moda estimada es siempre uno de los valores más probables. Esta observación ha resultado ser un comportamiento general para cada s y cada T_1 . Por tanto, la estimación de la moda puede considerarse fiable.

s	11	13	15
Valores (%)	9 (11 %)	11 (9 %)	13 (3 %)
	10 (9 %)	12 (6 %)	14 (10 %)
	11 (18 %)	13 (12 %)	15 (18 %)
	12 (17 %)	14 (17 %)	16 (16 %)
	13 (15 %)	15 (18 %)	17 (12 %)
	14 (13 %)	16 (8 %)	18 (7 %)
	15 (8 %)	17 (7 %)	19 (14 %)

Tabla 4.3: Simulación PC 1. Valores más frecuentes del número de puntos antes de s y sus frecuencias del proceso $N(t, B)$ para $T_1 = 10$.

Por último, pasemos a examinar el comportamiento de las estimaciones de la probabilidad de tener una nueva ocurrencia antes de un instante de tiempo s . La figura 4.1 presenta un gráfico de cajas y bigotes de las estimaciones de las citadas probabilidades dadas en los instantes $T_1 = 10$ y $s = 11, 15$ y 20 , que fueron los mismos que se utilizaron para medir el error de la estimación de

la moda, y en los que se puede ver la precisión de las estimaciones obtenidas para la probabilidad de tener una nueva ocurrencia antes de un instante de tiempo s .

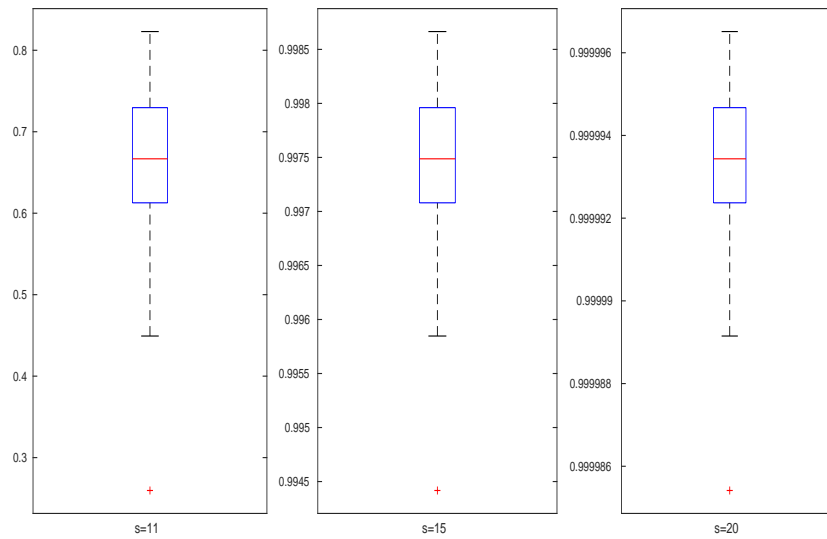


Figura 4.1: Simulación PC 1. Gráfico de las estimaciones de la probabilidad de tener una nueva ocurrencia antes de un tiempo s para $T_1 = 10$ del proceso $N(t, B)$.

4.2.2. Simulación PC caso 2

Para la segunda simulación, se ha considerado el caso de un PC compuesto que puede ser interpretado como mezcla de procesos de Poisson con ocurrencias simultáneas, donde las marcas en un instante de tiempo, es el número de ocurrencias en ese momento.

En concreto, se ha tomado un PC compuesto que a su vez está formado por cuatro procesos de Poisson aleatoriamente simultáneos, y cuyas intensidades son 5, 4, 6 y 3, respectivamente. Cada uno de estos procesos sucede con probabilidad 0.5 en cada instante de tiempo.

En este caso, las marcas del proceso se distribuyen según una distribución Binomial(10, 0.6), y como subconjunto de marcas se ha escogido $B = \{u; 6 \leq u \leq 8\}$ de forma que calculando su probabilidad, se obtiene $\int_B P_u(dU) = 0.4658$.

En esta simulación, de modo análogo a la simulación anterior, se ha fijado el intervalo del futuro ($T_1 = 5, T_2 = 20$], y realizando los cálculos pertinentes resulta que el proceso media puede modelizarse como un $PCP(q_2; p_1, \dots, p_j) \equiv PCP(9; 4, 4, 0, 4, 4, 4, 4, 0, 4)$. Descartando aquellas componentes que no están explicadas por ninguna en el pasado, el modelo cuenta con siete componentes en este caso. El proceso media para el instante de tiempo $s \in (T_1 = 5, T_2 = 20]$ puede escribirse como

$$\tilde{\Lambda}^9(s) = \mu_{\Lambda}^2(s) + \begin{pmatrix} -0.696 & -4.007 & -3.017 & -2.367 \\ 1.876 & 1.099 & 1.427 & 0.650 \\ -0.988 & 0 & -0.365 & -0.327 \\ -0.194 & -0.387 & 0 & 0.004 \\ 0.021 & 0.134 & 0 & -0.134 \\ -0.594 & 0.093 & 0 & 0.107 \\ 0.019 & 0 & 0 & 0.364 \end{pmatrix} \times \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{pmatrix} (g_1(s), g_2(s), g_4(s), g_5(s), g_6(s), g_7(s), g_9(s)).$$

Si el intervalo del futuro fuese distinto al escogido, podrían realizarse estos mismos cálculos para dicho intervalo resultando la ecuación correspondiente.

Los resultados obtenidos para las estimaciones realizadas en este ejemplo se muestran en la tabla 4.4. De nuevo, para una mejor ilustración, se han escogido tres intervalos del futuro diferentes, ($T_1 = 5, T_2 = 20$], ($T_1 = 10, T_2 =$

20] o $(T_1 = 15, T_2 = 20]$.

s	$T_1 = 5$			$T_1 = 10$			$T_1 = 15$		
	Media	Moda	Prob	Media	Moda	Prob	Media	Moda	Prob
6	31.882	31	32	0.983					
7	37.194	36	37	0.999					
8	42.420	41	42	0.999					
9	47.738	46	47	0.999					
10	52.989	52	53	0.999					
11	58.271	57	58	0.999	58.271	57	58	0.999	
12	63.578	62	63	1	63.578	62	63	0.999	
13	68.879	67	68	1	68.879	67	68	0.999	
14	74.226	73	74	1	74.226	73	74	0.999	
15	79.495	78	79	1	79.495	78	79	0.999	
16	84.770	83	84	1	84.770	83	84	0.999	84.770
17	90.085	89	90	1	90.085	89	90	1	90.085
18	95.436	94	95	1	95.436	94	95	1	95.436
19	100.715	99	100	1	100.715	99	100	1	100.715
20	105.875	104	105	1	105.875	104	105	1	105.875

Tabla 4.4: Simulación PC 2. Estimación de la media, moda y probabilidad de tener una nueva ocurrencia antes de un tiempo s (Prob) del proceso $N(t, B)$.

En la tabla 4.5 se muestran una comparativa entre las medias estimadas del proceso $N(s, B)$ y las medias muestrales obtenidas de cada una de las trayectorias simuladas.

De nuevo se observa la gran similitud de los datos obtenidos en ambos casos. Es de destacar que, como ya se ha comentado con anterioridad, cuanto más cerca se encuentra el instante s del instante T_1 mayor es dicha similitud.

Para hallar el número de puntos y los porcentajes muestrales antes del instante de tiempo s , de nuevo se han obtenido a partir de 100 nuevas trayectorias muestrales, siguiendo el patrón de la simulación anterior. Los datos resultantes se muestran en la tabla 4.6.

s	Estimación de la media			Medias muestrales
	$T_1 = 5$	$T_1 = 10$	$T_1 = 15$	
6	31.882			31.900
7	37.194			37.155
8	42.420			42.466
9	47.738			47.712
10	52.271			53.038
11		58.271		58.264
12		63.578		63.532
13		68.879		68.842
14		74.226		74.126
15		79.495		79.470
16		84.770		84.718
17		90.085		89.997
18		95.436		95.312
19		100.715		100.646
20		105.875		105.955

Tabla 4.5: Simulación PC 2. Comparativa de las medias estimadas con las medias muestrales del proceso $N(t, B)$ para cada trayectoria de $N(s, B)$.

$s = 10$		$s = 15$		$s = 20$	
Np	%	Np	%	Np	%
47	7	69	7	99	2
48	6	70	3	100	3
49	1	71	3	101	6
50	4	72	5	102	2
51	6	73	5	103	4
52	3	74	4	104	6
53	1	75	10	105	7
54	9	76	2	106	3
55	2	77	7	107	1
56	5	78	5	108	9
57	1	79	2	109	2
58	8	80	5	110	1

Tabla 4.6: Simulación PC 2. Valores más frecuentes del número de puntos antes de s y sus frecuencias del proceso $N(t, B)$ para $T_1 = 10$.

En el gráfico de cajas y bigotes 4.2 se observa la precisión de las estimaciones de la probabilidad de tener una nueva ocurrencia antes de un instante de tiempo s .

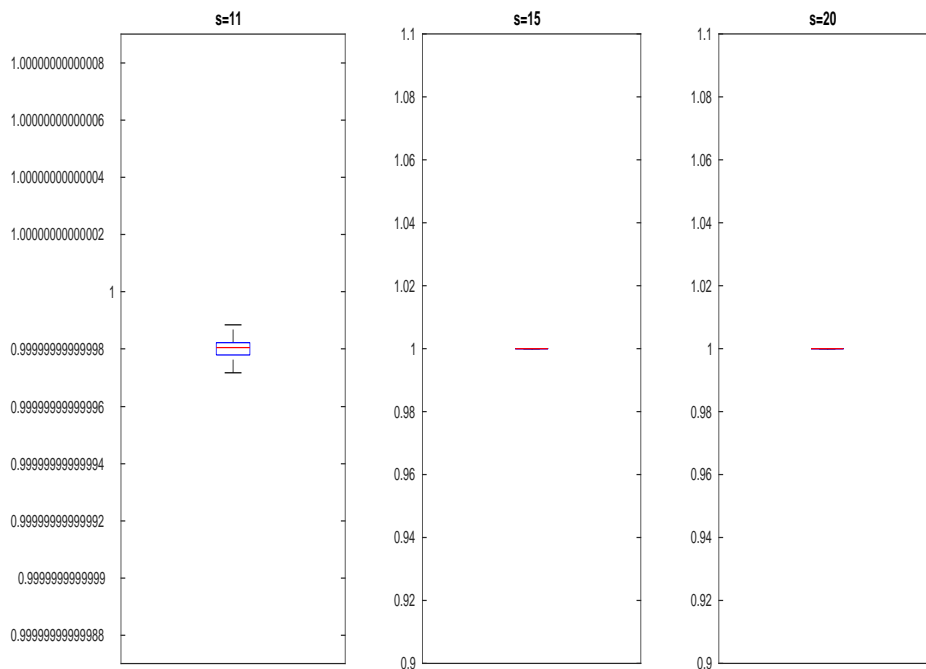


Figura 4.2: Simulación PC 2. Gráfico de las estimaciones de la probabilidad de tener una nueva ocurrencia antes de un tiempo s para $T_1 = 10$ del proceso $N(t, B)$.

4.2.3. Simulación PC caso 3

Retomamos ahora la simulación 1.6 que se introdujo en el Capítulo 1 como ilustración de los procesos con intensidad tipo fase.

Recordemos que esta simulación es relativa a un sistema de fiabilidad multiestado. Se ha asumido que el sistema está compuesto por tres estados, donde la distribución inicial, $\boldsymbol{\pi}$, determina el comportamiento hasta el fallo.

El tiempo de fallo de cada estado en cualquier tiempo está distribuido exponencialmente con tiempos de permanencia aleatorios, P_1 , P_2 y P_3 , respectivamente, y cada vez que el sistema falla, se produce una señal bidimensional independiente. Cuando el sistema falla, se cambia por uno nuevo idéntico al anterior. El tiempo hasta el fallo se distribuye mediante una distribución tipo fase con representación $PH(\boldsymbol{\pi}, T(\beta_1, \beta_2, \beta_3))$ donde

$$\boldsymbol{\pi} = (0.2, 0.4, 0.4) \text{ y } T = \begin{pmatrix} -1/\beta_1 & 0 & 0 \\ 0 & -1/\beta_2 & 0 \\ 0 & 0 & -1/\beta_3 \end{pmatrix}$$

con $\beta_1 \rightsquigarrow \mathcal{B}(10, 0.8) + 1$, $\beta_2 \rightsquigarrow \mathcal{B}(20, 0.5) + 1$ y $\beta_3 \rightsquigarrow \mathcal{B}(1, 0.5) + 1$.

Dado que en el citado ejemplo se consideraba un PC con intensidad tipo fase, ahora ampliaremos al caso en el que se tienen en cuenta las marcas del proceso. Consideraremos que las marcas son bidimensionales y distribuidas según la distribución Normal $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ con $\boldsymbol{\mu} = (4, 8)$ y $\boldsymbol{\Sigma} = \begin{pmatrix} 1.5 & 0.7 \\ 0.7 & 1.5 \end{pmatrix}$.

Como conjunto de marcas se ha escogido $B = \{(x_1, x_2); 2 \leq x_1 \leq 4, 7 \leq x_2 \leq 9\}$ con probabilidad, $\int_B P_u(dU) = 0.2717$. La figura 4.3 representa el espacio de marcas con su distribución y el espacio B .

Habiendo definido el proceso intensidad, el espacio de marcas y el conjuntos de marcas B , el PC compuesto se define como el PC formado por los puntos con marcas en B . La figura 4.4 muestra varias trayectorias muestrales del PC sin tener en cuenta las marcas y el PC con marcas en el conjunto B .

Siguiendo los mismos pasos que las simulaciones anteriores, el proceso media puede modelarse como un $PCP(9; 4, 4, 0, 0, 4, 0, 0, 0, 4)$ para $s \in (T_1 = 5, T_2 = 20]$. Teniendo en cuenta aquellas componentes que están correlacionadas con alguna del pasado, el modelo cuenta con cuatro componentes del

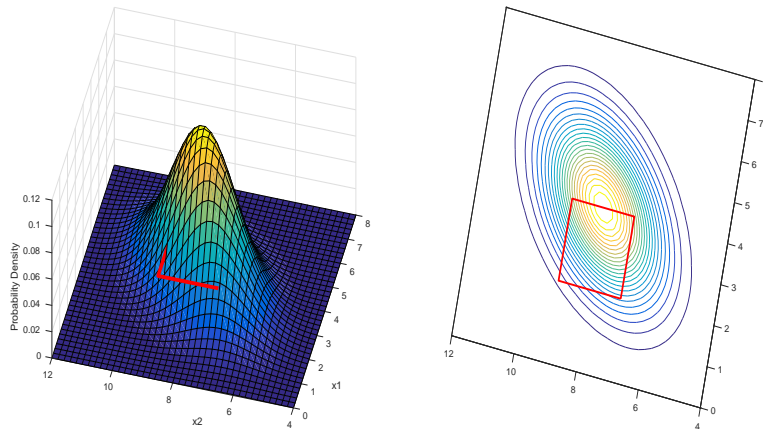


Figura 4.3: Simulación PC 3. Gráfico del espacio de marcas con su distribución y el conjunto de marcas B del proceso $N(t, B)$.

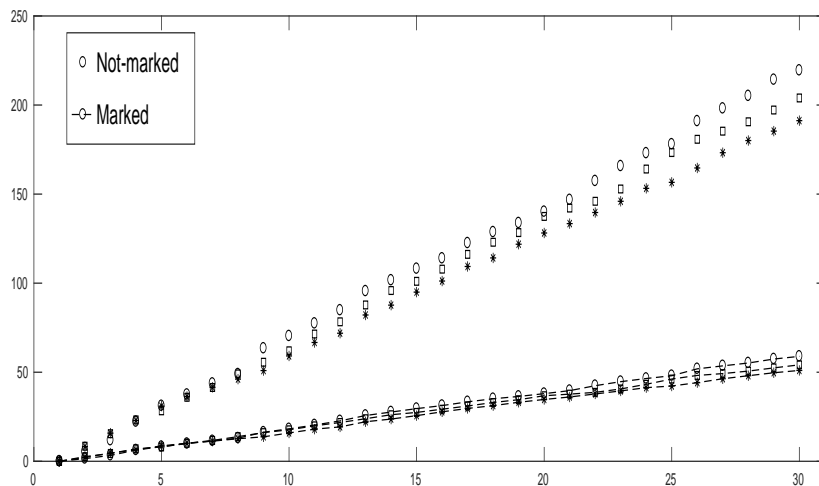


Figura 4.4: Simulación PC 3. Trayectorias muestrales del PC sin tener en cuenta las marcas y el PC con marcas en el conjunto B del proceso $N(t, B)$.

futuro y se expresa como

$$\tilde{\Lambda}^9(s) = \mu_{\Lambda}^{(2)}(s) + \begin{pmatrix} 0 & 0 & -0.449 & -0.525 \\ 0.615 & -0.501 & 0.073 & 0.350 \\ -0.013 & -0.123 & 0.161 & 0.312 \\ 0 & 0 & 0.262 & 0.251 \end{pmatrix} \times \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{pmatrix} (g_1(s), g_2(s), g_5(s), g_9(s)).$$

Este ejemplo puede interpretarse como un PC espacio-temporal en el que la marca de un punto representa su posición en el plano.

En la tabla 4.7 se muestran los resultados de las estimaciones obtenidas para esta simulación. Al igual que en los dos casos anteriores, las estimaciones se han realizado en tres intervalos del futuro distintos ($T_1 = 5, T_2 = 20$], ($T_1 = 10, T_2 = 20$] o ($T_1 = 15, T_2 = 20$], con el fin de una mejor ilustración de los mismos.

Pasemos a evaluar ahora los resultados obtenidos en la simulación 4.2.3 del mismo modo que se ha procedido en las simulaciones anteriores. Los datos de las medias muestrales obtenidos para el PC compuesto con intensidad tipo fase descrito anteriormente pueden verse en la tabla 4.8, y al comparar con los datos obtenidos en la tabla 4.7 se observa la similitud de los mismos.

Los valores más frecuentes del número de puntos antes de un instante de tiempo s así como sus frecuencias, se han obtenido a partir de la simulación de 100 nuevas trayectorias muestrales, los resultados son los expuestos en la tabla 4.9

Examinando ahora las estimaciones de la probabilidad de tener una nueva ocurrencia antes de un instante de tiempo s , podemos observar en el

s	$T_1 = 5$			$T_1 = 10$			$T_1 = 15$		
	Media	Moda	Prob	Media	Moda	Prob	Media	Moda	Prob
6	11.041	10	11	0.831					
7	12.941	11	12	0.977					
8	14.769	13	14	0.996					
9	16.629	15	16	0.999					
10	18.452	17	18	0.999					
11	20.265	19	20	0.999	20.265	19	20	0.583	
12	22.183	21	22	0.999	22.183	21	22	0.972	
13	24.085	23	24	0.999	24.085	23	24	0.995	
14	25.911	24	25	0.999	25.911	24	25	0.999	
15	27.751	26	27	0.999	27.751	26	27	0.999	
16	29.559	28	29	0.999	29.559	28	29	0.999	29.559
17	31.259	30	31	0.999	31.359	30	31	0.999	31.359
18	33.285	32	33	0.999	33.285	32	33	0.999	33.285
19	35.166	34	35	0.999	35.166	34	35	0.999	35.166
20	36.936	35	36	0.999	36.936	35	36	0.999	36.936

Tabla 4.7: Simulación PC 3. Estimación de la media, moda y probabilidad de tener una nueva ocurrencia antes de un tiempo s (Prob) del proceso $N(t, B)$.

s	Estimación de la media			Medias muestrales
	$T_1 = 5$	$T_1 = 10$	$T_1 = 15$	
6	11.041			11.147
7	12.941			12.993
8	14.769			14.891
9	16.629			16.716
10	18.452			18.588
11		20.265		20.414
12		22.183		22.225
13		24.085		24.118
14		25.919		26.016
15		27.751		27.843
16		29.559		29.711
17		31.359		31.528
18		33.285		33.308
19		35.166		35.246
20		39.936		37.112

Tabla 4.8: Simulación PC 3. Comparativa de las medias estimadas y las medias muestrales para cada trayectoria de $N(s, B)$ del proceso $N(t, B)$.

$s = 10$		$s = 15$		$s = 20$	
Np	%	Np	%	Np	%
7	1	14	6	26	8
8	5	15	6	27	5
9	4	16	4	28	3
10	5	17	7	29	5
11	10	18	6	30	11
12	7	19	8	31	4
13	8	20	6	32	8
14	11	21	5	33	4
15	8	22	3	34	3
16	7	23	7	35	5
17	10	24	4	36	6
18	5	25	9	37	3
19	3	26	4	38	4

Tabla 4.9: Simulación PC 3. Valores más frecuentes del número de puntos antes de s y sus frecuencias del proceso $N(t, B)$ para $T_1 = 10$.

correspondiente gráfico de cajas y bigotes de la figura 4.5 la precisión de las estimaciones.

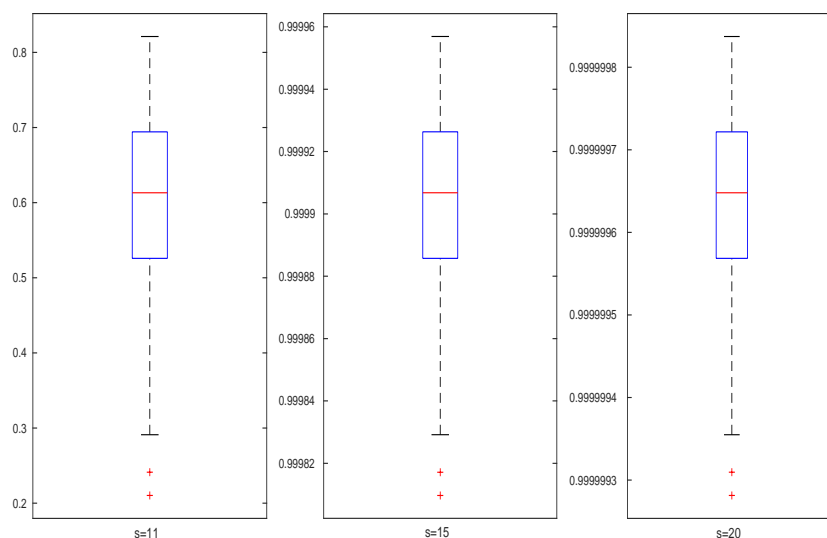


Figura 4.5: Simulación PC 3. Gráfico de las estimaciones de la probabilidad de tener una nueva ocurrencia antes de un tiempo s para $T_1 = 10$ del proceso $N(t, B)$.

Evaluación del test de bondad de ajuste

En este epígrafe, se seguirá tabajando con el mismo caso de simulación pero ahora con el objetivo de ilustrar el comportamiento del test de bondad de ajuste propuesto para un PC compuesto.

Antes de entrar en ello, merece la pena comentar la siguiente anécdota. Este contraste, como se explicó en la Sección 1.4.3, fue inicialmente inspirado en el recuento de emisiones de partículas por isótopos radiactivos. En ese caso, durante una de las pruebas de laboratorio, ocurrió que bajó la intensidad de corriente de forma momentánea sin que el técnico de laboratorio pudiera notarlo. El contraste de bondad de ajuste rechazó la trayectoria, lo

que resultó sorprendente. Después de las comprobaciones, se supo acerca de la bajada de tensión. En conclusión, el test fue capaz de detectar un incidente tan pequeño y supuso un respaldo en la confianza de que la evaluación del test iba a ser satisfactoria como así ocurrió (ver Bouzas et al., 2010b [26]).

El caso que se estudia en este momento corresponde al de un ejemplo del contraste cuya hipótesis nula es un proceso con intensidad conocida. De hecho, se sabe que la media y la varianza del proceso tipo fase que se usa como intensidad y descrito en la ecuación (1.6) son las siguientes

$$\mu(t) = -\pi T^{-1} \mathbf{1} \quad \sigma^2(t) = 2\pi T^{-2} \mathbf{1} - (\pi T^{-1} \mathbf{1})^2$$

Se ha simulado 100 veces el proceso para obtener 100 nuevas trayectorias de su media, $\mu(t)$. La figura 4.6 muestra estas trayectorias, y la figura 4.7 muestra solo 3 de ellas para apreciarlas mejor.

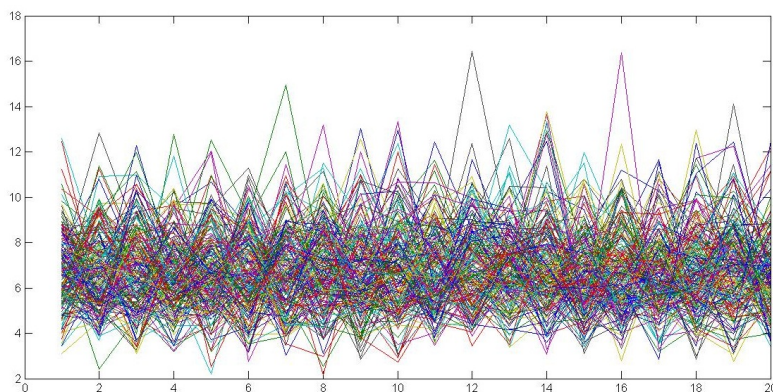


Figura 4.6: Simulación PC 3. 100 nuevas trayectorias de la media del proceso intensidad.

Para observar el comportamiento del test, se han llevado a cabo dos actuaciones diferentes. En primer lugar, se considera el PC sin marcas y se corrompen las nueva trayectorias experimentales, μ_{exp} , eliminando puntos de las mismas a razón de M/c , donde M es la media de puntos en ese instante en las 100 trayectorias y $c = 3, 2.75, 2.5, 2.25$ y 2 . Aplicando el test para

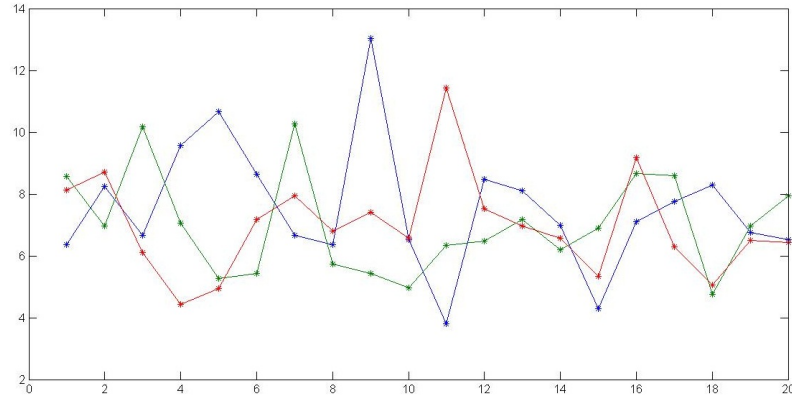


Figura 4.7: Simulación PC 3. 3 de las nuevas trayectorias de la media del proceso intensidad.

comprobar si las trayectorias corresponden al proceso cuya intensidad se define en (1.6) a las trayectorias sin corromper (intactas) o con la corrupción comentada, el porcentaje de trayectorias no rechazadas se pueden observar en la tabla 4.10.

μ_{exp}	Tray.intactas	c=3	c=2.75	c=2.5	c=2.25	c=2
%	87	64	46	31	17	6

Tabla 4.10: Simulación PC 3. Porcentaje de trayectorias no rechazadas por el contraste de bondad de ajuste sin corromper o eliminando cada vez más puntos.

Se observa, como era de esperar, que cuantos más puntos se han eliminado, menos trayectorias son consideradas pertenecientes al proceso cuya intensidad tiene la media de la hipótesis nula y, por tanto, no siguen el patrón supuesto.

La otra actuación llevada a cabo consiste en considerar el PC compuesto del ejemplo de simulación de esta sección, con marcas distribuidas como ya se propuso, distribución Normal $\mathcal{N}(\mu, \Sigma)$ con $\mu = (4, 8)$ y $\Sigma = \begin{pmatrix} 1.5 & 0.7 \\ 0.7 & 1.5 \end{pmatrix}$, y

considerando el mismo subconjunto $B = \{(x_1, x_2); 2 \leq x_1 \leq 4, 7 \leq x_2 \leq 9\}$ con probabilidad $p = 0.2717$. A continuación, se considerará dicho proceso como proceso observado y se tomarán como diferentes patrones varios procesos. Estos procesos patrones se han escogido de forma que la probabilidad (p) de que un punto pertenezca al proceso es menor o mayor que la de pertenecer al observado ($p = 0.2717$). Los porcentajes de trayectorias no rechazadas se pueden observar en la tabla 4.11.

<i>Proceso Observado</i>	<i>Patrón</i>									
	$p = 0.1$	$p = 0.15$	$p = 0.2$	$p = 0.25$	$p = 0.2717$	$p = 0.3$	$p = 0.35$	$p = 0.4$	$p = 0.45$	$p = 0.5$
$N(t, B) p = 0.2717$										
Porcentaje (%)	0	10	68	88	89	86	86	72	71	57

Tabla 4.11: Simulación PC 3. Porcentaje de trayectorias no rechazadas por cada contraste de bondad de ajuste según el valor de la probabilidad p .

En esta segunda actuación, se observa un comportamiento similar al anterior. Cuanto más se aleja la probabilidad del valor 0.2717, menos trayectorias son consideradas pertenecientes al proceso patrón supuesto. Por tanto, el comportamiento del contraste sigue siendo adecuado.

4.3. Aplicaciones del PC compuesto

En esta sección queremos centrar la atención en la utilidad que tienen los PC y PC compuestos a la hora de modelizar casos reales. Por este motivo, los resultados teóricos expuestos a lo largo de la memoria, no solo han sido utilizados en simulaciones, tal y como se ha mostrado en la sección anterior, sino que se han aplicado a casos con datos reales. Los resultados obtenidos han sido muy satisfactorios, como puede verse en cada una de las aplicaciones siguientes.

4.3.1. Aplicación 1: Puntos de cambio de tendencia

El primer caso con datos reales al que aplicaremos ahora los resultados teóricos, son los cambios de tendencia en las cotizaciones bursátiles, para ello, consideraremos estos cambios como los sucesos de un PC. Dado que el precio de las acciones del mercado de valores depende de las cotizaciones anteriores, pero no necesariamente los cambios de tendencia, se aplica la metodología a diferentes acciones y también al conjunto de ellas para probar su robustez. De hecho, como veremos más adelante, se estiman bastante bien los futuros puntos de cambio de tendencia. Los datos que se han utilizado corresponden a las bolsas de Atenas y Nueva York.

Estimación de cambios de tendencia

En primer lugar, presentamos un ejemplo de la estimación de cambios de tendencia en las cotizaciones de un único tipo de acciones. Se dispone de 23 trayectorias muestrales obtenidas de 200 observaciones diarias consecutivas del precio de la acción al cierre del mercado.

Cortando en subtrayectorias para aplicar la estimación del proceso media, el intervalo de estudio se ha escogido $[T_0, T_2] = [0, 20]$. En el caso en el que $T_1 = 15$, el proceso media puede estimarse del decimosexto hasta el decimonoveno, a partir del conocimiento que tenemos de los 15 primeros días. En este caso se obtiene que la ecuación de estimación del proceso media dada en (1.29) sería

$$\tilde{\Lambda}^{q_2=1}(s) = \mu_{\Lambda}^{(2)}(s) + \sum_{j=1}^1 \left(\sum_{i=1}^{p_j} b_i^j \xi_i \right) g_j(s); \quad s \in (T_1, T_2)$$

con $p_j = 1, 2, 6$, para $j = 1 = q_2$, respectivamente. Por tanto, el cálculo de $\tilde{\Lambda}^1(s)$, $s \in (T_1, T_2)$ permite la estimación de la media y la moda del número de cambios de tendencia en cualquier día del intervalo del futuro dentro del intervalo general $[0, 20)$ cuando se observa una nueva trayectoria muestral.

Para probar la precisión de las estimaciones se reserva una de las trayectorias muestrales y se trata como un nuevo dato. Aplicando la metodología expuesta, se calculan al igual que la media, la media de cambios de tendencia hasta un tiempo s , y la moda. Para la media, los valores estimados y los valores reales en el intervalo del futuro se presentan en la tabla 4.12, donde ε es la raíz del error cuadrático medio (RECM) de la estimación. Se debe señalar que los valores probables de $N(t)$ son $n = 0, 1, 2, \dots$ por lo que la probabilidad correspondiente a cada uno de ellos puede ser muy pequeña. Así, una estimación precisa de la moda es muy difícil. Sin embargo, en los cálculos se ha observado que cuando se estima la moda, aunque ésta no sea exactamente la misma que la real, suele diferir en no más de una unidad. Por este motivo, se ha incluido el intervalo intercuartílico de la moda, $I(n_{\text{máx}}(s))$, y entonces se ha llamado precisión al porcentaje de estimaciones en ese intervalo.

Estos mismos cálculos, pueden realizarse para cualquier valor de T_1 . En la tabla 4.13 se muestra un resumen de los cálculos obtenidos para varios

s	$E[N(s)]$	$\hat{E}[N(s)]$	$n_{\text{máx}}(s)$	$\hat{n}_{\text{máx}}(s)$	$I(n_{\text{máx}}(s))$
16	7.6	8.2636	8	8	[6, 8.75]
17	8.3	8.7182	9	8	[7, 9]
18	8.8	9.2182	9	9	[6, 9]
19	9.3	9.7818	9	9	[7.25, 9]
$\varepsilon = 0.5055$			precisión = 100 %		

Tabla 4.12: Aplicación 1. Media y moda para diferentes valores de s para $T_1 = 15$, error y ajuste.

valores de T_1 , lo que implica estimar distintos intervalos del futuro a partir de sus respectivos pasados.

Con el fin de seguir comprobando la precisión de la metodología, y puesto que se tienen 23 trayectorias muestrales, el método fué aplicado a los datos 23 veces, dejando fuera una trayectoria distinta cada vez, para que actuara como nuevo dato. Entonces, se obtiene el RECM y el ajuste para cada trayectoria muestral reservada y cada valor fijo de T_1 . La tabla 4.14 y la figura 4.8 muestran la media y la desviación media de los errores para cada T_1 y una gráfica de ellos.

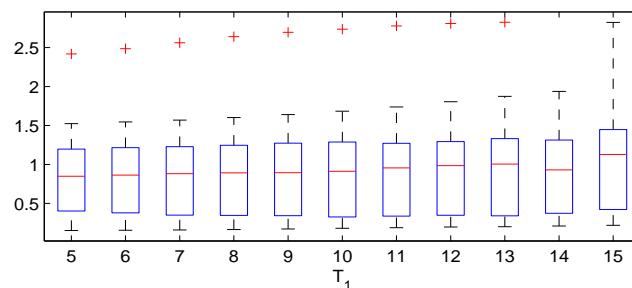


Figura 4.8: Aplicación 1. Gráfico de cajas y bigotes del RECM de la estimación de la media con todas las trayectorias muestrales para distintos valores de T_1 .

	s	$E[N(s)]$	$\hat{E}[N(s)]$	$n_{\text{máx}}(s)$	$\hat{n}_{\text{máx}}(s)$	$I(n_{\text{máx}}(s))$
$T_1 = 5$	6	3	3.1136	4	3	[2, 3]
	7	3.8	3.6045	4	3	[2.25, 3.75]
	8	4.3	4.1273	4	4	[3, 4]
	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
	18	8.8	9.2182	9	9	[6, 9]
	19	9.3	9.7818	9	9	[7.25, 9]
		$\varepsilon = 0.3048$		precisión = 100 %		
$T_1 = 8$	9	4.8	4.6636	5	4	[3, 5]
	10	5.3	5.1636	5	5	[3, 5]
	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
	14	7.1	7.2727	7	7	[5, 7]
	18	8.8	9.2182	9	9	[6, 9]
	19	9.3	9.7818	9	9	[7.25, 9]
		$\varepsilon = 0.333$		precisión = 100 %		
$T_1 = 10$	11	5.9	5.7409	6	5	[4, 6]
	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
	18	8.8	9.2182	9	9	[6, 9]
	19	9.3	9.7818	9	9	[7.25, 9]
		$\varepsilon = 0.3624$		precisión = 100 %		

Tabla 4.13: Aplicación 1. Media y moda para diferentes valores de s y T_1 , error y ajuste.

	T_1					
	5	6	7	8	9	10
$\bar{\varepsilon}$	0.8490	0.8553	0.8667	0.8804	0.8942	0.9105
$\sigma(\varepsilon)$	0.5437	0.5636	0.5840	0.6019	0.6181	0.6321
	T_1					
	11	12	13	14	15	
$\bar{\varepsilon}$	0.9291	0.9507	0.9733	0.9139	1.0464	
$\sigma(\varepsilon)$	0.6443	0.6564	0.6683	0.5548	0.6763	

Tabla 4.14: Aplicación 1. Media y desviación típica del RECM (ε) de las estimaciones medias del número de cambios de tendencia a partir de todas las trayectorias muestrales.

Aplicación del test de bondad de ajuste

Una vez estudiado el caso de los cambios de tendencia en un único tipo de acciones, pasemos ahora al estudio del comportamiento de los cambios de tendencia de varios tipos de acciones distintas, que resulta ser similar al anterior. Con el fin de evaluar esta similitud, se aplica un test de bondad de ajuste en el que eligiendo el proceso de recuento de los cambios de tendencia de un mercado, y tomándolo como patrón, el test de bondad de ajuste evalúa si los cambios de tendencia de cualquier otra acción siguen este mismo modelo. Reiterando este procedimiento con cada una de las acciones disponibles, el test dió como resultado que en ningún caso podía descartarse dicha similitud. Por tanto, los datos disponibles de cada acción, han sido tratados como una trayectoria muestral del mismo fenómeno. Este comportamiento se puede ver en la figura 4.9. Las trayectorias muestrales provienen de las acciones (ALPHA, CAT, EK, ELL, etc.) en las bolsas de Atenas y Nueva York como se explicó anteriormente.

Dada la conclusión del test, se puede volver la estimación de los estadísti-

cos juntando ambos mercados de valores. Se han observado 10 trayectorias muestrales durante 3600 días consecutivos. Considerando que todas las trayectorias muestrales provienen del mismo PC, se han estimado la media y la moda de los cambios de tendencia de un tipo de acción general en un intervalo del futuro para cualquier periodo de 50 días, es decir $[0, 50]$. Aplicando el PCP para un T_1 dado, se obtiene la estimación del proceso media, y a partir de ésta, la estimación de la media y la moda para cada instante $s \in (T_1, T_2)$. Estas estimaciones podrían calcularse para distintos valores de T_1 , como ya se ha mencionado. En esta ocasión se ofrece un resumen de ellos en la tabla 4.15 así como su correspondiente gráfico de cajas y bigotes en el que se muestra la precisión de la estimación, en la figura 4.10.

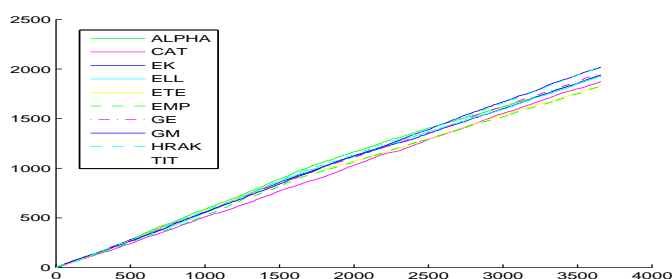


Figura 4.9: Aplicación 1. Trayectorias muestrales de los cambios de tendencia de las acciones ALPHA, CAT, EK, etc.

4.3.2. Aplicación 2: Datos extremos en meteorología

Las condiciones climáticas suelen ser sujeto de estudio por su enorme impacto socioeconómico y ambiental, por lo que el estudio de métodos para su comparación y predicción ha suscitado creciente interés en las últimas décadas. En particular, los valores de las variables climáticas que exceden

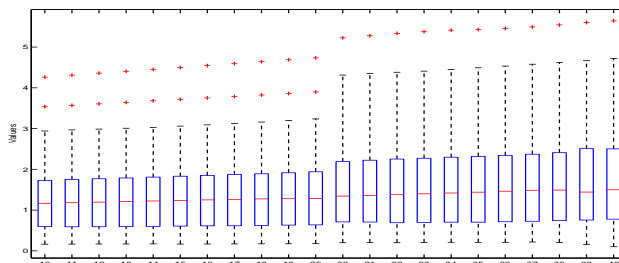


Figura 4.10: Aplicación 1. Gráfico de cajas y bigotes del RECM de la estimación de la media entre todas las trayectorias muestrales para distintos valores de T_1 .

	T_1							
	10	11	12	13	14	15	16	17
$\bar{\varepsilon}$	1.3298	1.3421	1.3551	1.3686	1.3825	1.3963	1.4102	1.4243
$\sigma(\varepsilon)$	0.9000	0.9109	0.9217	0.9324	0.9430	0.9530	0.9629	0.9729
	T_1							
	18	19	20	21	22	23	24	25
$\bar{\varepsilon}$	1.4384	1.4524	1.4658	1.4804	1.4947	1.5099	1.5246	1.5396
$\sigma(\varepsilon)$	0.9836	0.9947	1.0060	1.0177	1.0288	1.0399	1.0508	1.0619
	T_1							
	26	27	28	29	30	31	32	33
$\bar{\varepsilon}$	1.5551	1.5699	1.5851	1.6007	1.6166	1.6330	1.6485	1.6641
$\sigma(\varepsilon)$	1.0732	1.0861	1.0997	1.1139	1.1274	1.1401	1.1524	1.1652
	T_1							
	34	35	36	37	38	39	40	
$\bar{\varepsilon}$	1.6806	1.6954	1.7116	1.7279	1.7440	1.7547	1.7735	
$\sigma(\varepsilon)$	1.1792	1.1922	1.2031	1.2159	1.2284	1.2534	1.2583	

Tabla 4.15: Aplicación 1. Media y desviación típica del RECM (ε) de las estimaciones de la media para el número de cambios de tendencia entre todas las trayectorias muestrales.

ciertos límites, o valores extremos, tienen una importancia relevante. Dado que las ocurrencias de estos valores extremos conforman un proceso puntual, hemos considerado interesante aplicar la metodología desarrollada en esta memoria a dichos valores.

Los datos disponibles para esta aplicación son las temperaturas máximas y mínimas de estaciones meteorológicas de una misma zona climática en el sur de España. En concreto, se trata de 15 estaciones meteorológicas de las regiones de Andalucía, Badajoz, Murcia, y las ciudades de Ceuta y Melilla, para el periodo comprendido entre 1991 y 2011.

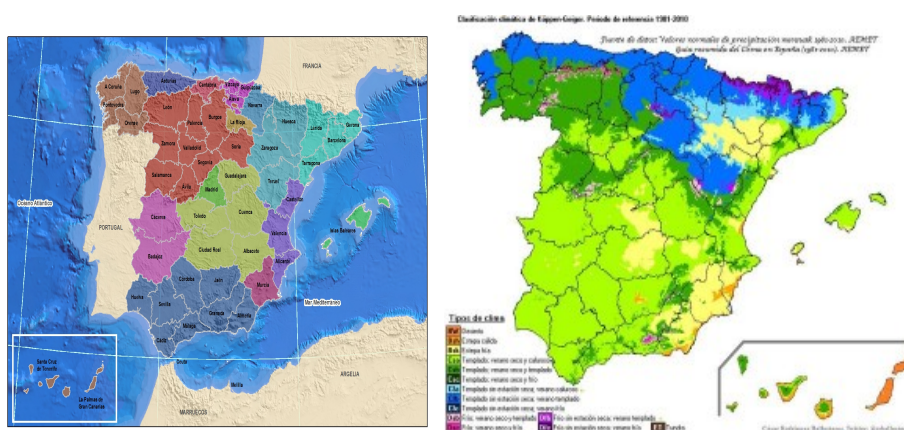


Figura 4.11: Aplicación 2. Provincias y zonas climáticas de España.

Se consideran como valores extremos, tanto para las temperaturas máximas como para las mínimas, aquellos que se encuentran fuera de un intervalo del 80 % alrededor de la media muestral en cada instante de observación, utilizando la desigualdad de Chebyshev. El intervalo resultante tras establecer esta condición, es $(\bar{x} - 2.361 \sigma, \bar{x} + 2.361 \sigma)$. Los valores extremos de máximas o mínimas conforman sendos procesos puntuales, cuyos recuentos se modelizan mediante un proceso de Cox y, como tales, están caracterizados mediante su proceso intensidad o su proceso media.

Tras la evaluación de los extremos, se puede observar que existen algu-

nos valores que son extremadamente altos o bajos; éstos pueden haber sido ocasionados por circunstancias inusuales. Los que resultan de interés son los puntos que representan el comportamiento habitual de los valores extremos de temperatura en la zona, y, por tanto, formarán los puntos pertenecientes al subconjunto B . Formalmente, el subconjunto B es la unión de los intervalos $(\bar{x} + 2.361\sigma, \bar{x} + 2.5\sigma)$ and $(\bar{x} - 2.361\sigma, \bar{x} - 2.5\sigma)$, donde, como ya fue comentado con anterioridad, 2.361 es el valor dado por la desigualdad de Chebyshev y 2.5 ha sido elegido con el fin de determinar un límite adicional a las marcas, tanto superior como inferior.

Estimación de valores máximos y mínimos

En primer lugar, se presentará detalladamente el estudio realizado para las temperaturas máximas; luego, el caso de las mínimas podrá exponerse de forma concisa.

En este primer caso de las temperaturas máximas, el subconjunto de marcas B tiene como probabilidad asociada $\int_B P_u(dU) = 0.5752$. Para ilustrar la selección de valores extremos y el subconjunto de marcas B , se ha escogido como ejemplo la provincia de Sevilla. Los valores de las temperaturas máximas diarias y los valores extremos, así como aquellos que pertenecen al subconjunto B pueden verse en la figura 4.12. Ampliando para una mejor visualización, los datos de la primera parte del año se pueden observar en la gráfica 4.13.

Realizando este procedimiento de selección de los datos en B para cada una de las estaciones meteorológicas, y teniendo en cuenta el proceso de recuento que se genera, $N(t, B)$, se obtienen 15 trayectorias muestrales. La figura 4.14 muestra las trayectorias muestrales de los procesos $N(t)$ y $N(t, B)$ para todas las estaciones meteorológicas de la aplicación.

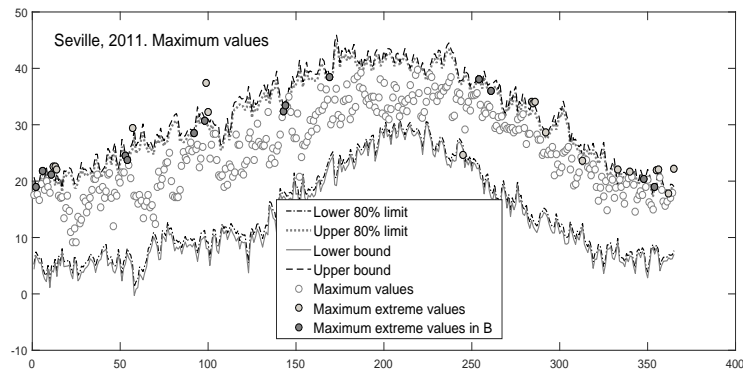


Figura 4.12: Aplicación 2. Temperaturas máximas diarias, extremos máximos y extremos máximos en B en la estación meteorológica de Sevilla en 2011.

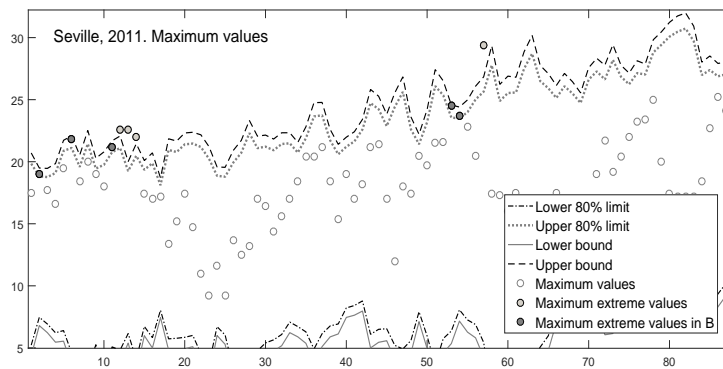


Figura 4.13: Aplicación 2. Temperaturas máximas diarias, extremos máximos y extremos en B de la estación meteorológica de Sevilla durante los primeros 80 días del 2011.

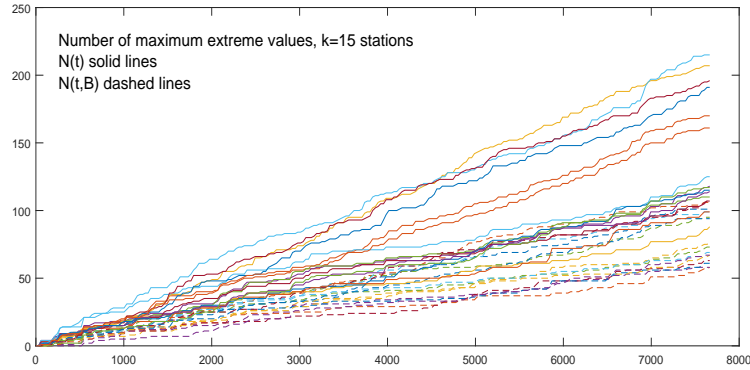


Figura 4.14: Aplicación 2. Trayectorias muestrales del proceso de recuento de los extremos de temperaturas máximas y los extremos en el subconjunto B de todas las estaciones meteorológicas.

En lo referente a los intervalos de tiempo utilizados en el ACP funcional y los posteriores resultados obtenidos, el intervalo $[T_0, T_2]$ es un año denotado por $[T_0 = 1, T_2 = 12]$, donde $1, 2, \dots$ son los correspondientes meses del año. Al igual que en los ejemplos anteriores, para la estimación de los parámetros del proceso, es necesario estimar previamente el proceso media que lo caracteriza. Dado que las temperaturas se han observado a lo largo de 21 años, una trayectoria muestral del proceso media puede estimarse a partir de 21 subtrayectorias del recuento de cada una de las 15 estaciones meteorológicas que forman parte del estudio. Es decir, se obtienen $k = 15$ trayectorias muestrales independientes, que dividimos en $r = 21$ trayectorias más cortas con $p = 12$ puntos de observación, con el fin de obtener puntos de estimación del proceso media.

Con el fin de estimar el proceso media, se utiliza ACP funcional que conserva la monotonía que da lugar a la ecuación (1.29). Habiendo realizado los cálculos, la modelización del proceso media de los valores extremos máximos resulta ser un PCP(1;2) para $s \in (T_1 = 5, T_2 = 12]$, puede escribirse

como:

$$\tilde{\Lambda}^1(s) = \mu_{\Lambda}^2(s) + \begin{pmatrix} -0.705 & 1.9792 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} g_1(s).$$

Una vez realizada la estimación de la media, podemos estimar la media del número de valores extremos, la moda y la probabilidad de tener un nuevo valor extremo antes de un tiempo dado. En la tabla 4.16 se muestran los resultados escogiendo, a modo de ejemplo, tres intervalos del futuro. Dichos intervalos comienzan en $T_1 = 4$ correspondiente con el mes de abril, 5 al mes de mayo y 6 al mes de junio, y los instantes de predicción son los siguientes meses.

El hecho de que se trabaje con datos reales no permite una evaluación de los resultado; no obstante, la comparación entre las estimaciones de la media y los valores de la media muestral puede ayudar a evaluar de algún modo la precisión de los resultados obtenidos. Con este fin y para una mejor ilustración, dicha comparativa se ha realizado en los tres escenarios de intervalos pasados y futuros escogidos anteriormente. Los resultados se muestran en la tabla 4.17, en la que puede observarse que las diferencias existentes entre los valores de las estimaciones y los muestrales son muy pequeñas.

Centrando la atención en los valores extremos para las temperaturas mínimas, los pasos a seguir serían los mismos que en el caso de las máximas, incluso los límites que se han utilizado y el subconjunto de marcas B escogido, son similares. De igual modo, la provincia de Sevilla sirve para ilustrar mejor la selección hecha de los valores extremos de temperaturas mínimas y de aquellos que se encuentran en B . Las temperaturas mínimas en Sevilla durante el año 2011 son las mostradas en la figura 4.15.

Para una mejor visualización de los datos, se ha generado la gráfica 4.16 en la cual se representan los datos de las temperaturas mínimas, sus extremos

T_1	s	Media	Moda		Prob
4	5	1.561	0	1	0.226
	6	1.876	0	1	0.435
	7	2.221	1	2	0.599
	8	2.473	1	2	0.689
	9	2.8	1	2	0.776
	10	3.077	2	3	0.830
11	3.455	2	3	0.883	
5	6	1.876	1	1	0.270
	7	2.221	2	1	0.483
	8	2.473	2	3	0.598
	9	2.8	2	3	0.710
	10	3.077	3	3	0.780
	11	3.455	3	3	0.843
6	7	2.221	2	1	0.088
	8	2.473	2	3	0.438
	9	2.8	2	3	0.602
	10	3.077	3	3	0.699
	11	3.455	3	3	0.794

Tabla 4.16: Aplicación 2. Estimación de la media, moda y probabilidad de tener un nuevo punto antes de un tiempo s para los valores extremos de temperatura máxima en el subconjunto B.

s	Estimación de la media			Media muestral
	$T_1 = 4$	$T_1 = 5$	$T_1 = 6$	
5	1.561			1.524
6		1.876		1.857
7			2.220	2.191
8			2.473	2.667
9			2.8	3.191
10			3.077	3.667
11			3.455	4

Tabla 4.17: Aplicación 2. Comparativa de las medias estimadas y las medias muestrales de $N(s, B)$ para los extremos de temperatura máxima en B.

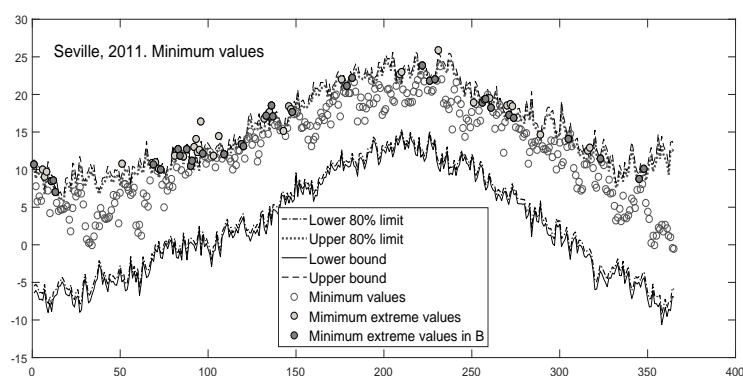


Figura 4.15: Aplicación 2. Temperaturas mínimas diarias, extremos mínimos y extremos en B de la provincia de Sevilla en 2011.

y sus extremos en B durante los primeros 80 días del 2011.

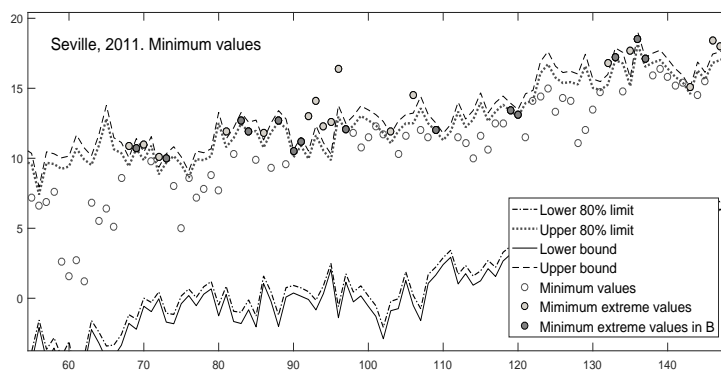


Figura 4.16: Aplicación 2. Temperaturas mínimas diarias, extremos mínimos y extremos en B de la provincia de Sevilla durante los primeros 80 días del 2011.

La figura 4.17 muestra la gráfica de las trayectorias muestrales de $N(t)$ y $N(t, B)$ obtenida tras haber seleccionado los datos de cada una de las estaciones meteorológicas y tenido en cuenta los procesos puntuales asociados a cada uno de los puntos.

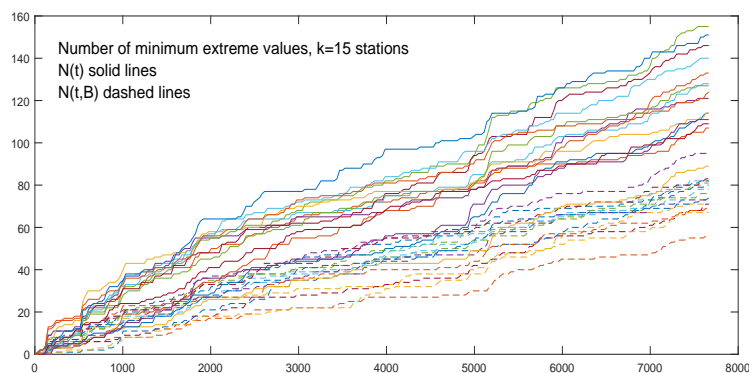


Figura 4.17: Aplicación 2. Trayectorias muestrales del proceso de recuento de los extremos de temperatura mínima y los extremos en el subconjunto B de todas las estaciones meteorológicas.

Haciendo uso de las trayectorias muestrales, se puede modelar el pro-

ceso media de los valores extremos de las temperaturas mínimas como un $PCP(2; 1, 1)$ para $s \in (T_1 = 5, T_2 = 11]$ y que tiene como representación:

$$\tilde{\Lambda}^2(s) = \mu_{\Lambda}^2(s) + \begin{pmatrix} 0.2277\xi_1 \\ -1.3651\xi_1 \end{pmatrix} (g_1(s), g_2(s)).$$

Una vez estimado el proceso media, pueden estimarse los distintos estadísticos de los valores extremos de las temperaturas mínimas en el subconjunto de marcas B , tal y como se muestran, para tres escenarios distintos, $T_1 =$ abril, mayo y junio, en la tabla 4.18.

Análogamente al caso de los extremos máximos, para poder valorar la precisión de los datos, dado que una evaluación no es posible al tratarse de datos reales, se ha realizado nuevamente una comparativa entre la estimación de las medias y los valores de las medias muestrales. Tal y como muestra la tabla 4.19, las diferencias son pequeñas.

Aplicación del test de bondad de ajuste

Se usará este conjunto de datos reales, no solo para estimar los estadísticos de los valores extremos si no para mostrar una aplicación del test de bondad de ajuste propuesto en esta memoria.

En primer lugar, se ha aplicado el test de bondad de ajuste sobre el proceso de recuento que modeliza las temperaturas mínimas, comprobando si cada una de las provincias del ámbito estudiado pertenece al modelo propuesto. La tabla 4.20 muestra los p-valores del criterio de Benjamini y Hochberg en la primera fila (α_j/p) y, a continuación, los correspondientes p-valores de todas las provincias. Se comprueba que solo Córdoba no cumple el requisito para asumir que sigue el mismo modelo (todos los p-valores mayores que los del criterio).

T_1	s	Media	Moda	Prob
4	5	1.417	0 1	0.265
	6	1.8	0 1	0.499
	7	2.223	1 2	0.674
	8	2.564	1 2	0.769
	9	2.855	1 2	0.828
	10	3.122	2 3	0.866
5	6	1.8	0 1	0.335
	7	2.223	1 2	0.577
	8	2.564	1 2	0.699
	9	2.855	1 2	0.775
	10	3.122	2 3	0.828
	11	3.361	2 3	0.864
6	7	2.223	1 2	0.372
	8	2.564	1 2	0.556
	9	2.855	1 2	0.669
	10	3.122	2 3	0.746
	11	3.361	2 3	0.788

Tabla 4.18: Aplicación 2. Estimación de la media, moda y probabilidad de tener un nuevo punto antes de s para los extremos de temperaturas mínimas en el subconjunto B.

s	Estimación de la media			Media muestral
	$T_1 = 4$	$T_1 = 5$	$T_1 = 6$	
5	1.417			1.571
6		1.800		1.714
7			2.223	2.143
8			2.564	2.381
9			2.856	2.714
10			3.122	2.905
11			3.361	3.191

Tabla 4.19: Aplicación 2. Comparativa de la estimación de la media y medias muestrales de $N(s, B)$ para las temperaturas mínimas.

t	α_j/p	0.0042	0.0083	0.0125	0.0167	0.0208	0.0250	0.0292	0.0333	0.0375	0.0417	0.0458	0.0500
p -valores	Badajoz	0.0047	0.0763	0.0906	0.3809	0.5012	0.5424	0.5778	0.6122	0.6304	0.6538	0.8641	0.9448
	Córdoba	0.0058	0.0059	0.0081	0.0356	0.1031	0.1992	0.2226	0.6087	0.6304	0.8003	0.8644	0.9448
	Granada	0.0113	0.1715	0.1914	0.2226	0.2370	0.2423	0.3419	0.4072	0.6541	0.8644	0.8993	0.9251
	Sevilla	0.0102	0.0211	0.0222	0.0281	0.0326	0.0340	0.1715	0.1718	0.3419	0.6954	0.9251	0.9375
	Morón	0.0465	0.1296	0.1622	0.1715	0.3104	0.4969	0.5296	0.6689	0.7979	0.8865	0.9375	0.9744
	Rota	0.0052	0.0364	0.1522	0.3586	0.3608	0.3638	0.3809	0.3979	0.5750	0.6142	0.6304	0.8973
	Jerez	0.0275	0.0652	0.0870	0.0927	0.1375	0.2423	0.3104	0.3275	0.3408	0.3608	0.9375	0.9756
	Cádiz	0.0642	0.1083	0.1544	0.2183	0.2826	0.3408	0.3809	0.5778	0.6304	0.8641	0.8644	0.8865
	Alcantarilla	0.0818	0.1296	0.1613	0.1781	0.3275	0.4037	0.6254	0.6304	0.6954	0.8641	0.9448	0.9475
	Huelva	0.0100	0.0397	0.0906	0.1201	0.1397	0.1585	0.1622	0.3015	0.4072	0.5296	0.6122	0.6689
	Granada A	0.0275	0.0740	0.0818	0.0967	0.1847	0.2183	0.2787	0.2826	0.3638	0.4979	0.6254	0.7077
	Melilla	0.0281	0.0870	0.1397	0.1732	0.3408	0.3608	0.6122	0.6538	0.6696	0.8641	0.9251	1.0000
	Málaga	0.0219	0.0465	0.0870	0.1481	0.1622	0.2370	0.3979	0.6087	0.6122	0.6689	0.7979	0.9475
	Almería	0.0131	0.0132	0.0218	0.0642	0.0692	0.0722	0.1613	0.1847	0.4072	0.6087	0.9375	0.9448
Murcia	0.0692	0.0763	0.1375	0.3499	0.5424	0.6122	0.7077	0.7324	0.8003	0.8865	0.9448	0.9744	

Tabla 4.20: Aplicación 2. P-valor para el test de bondad de ajuste sobre las temperaturas mínimas.

A continuación, se toman las provincias de Zaragoza y Orense, como ejemplos de fuera del área de estudio para ver si podrían considerarse similares o no. La tabla 4.21 muestra los p -valores correspondientes. Como consecuencia, se puede decir que para la provincia de Zaragoza se acepta que sigue el proceso propuesto y, por tanto, que sus extremos mínimos siguen un patrón similar a los de la zona meteorológica estudiada. Sin embargo, esto no es así para la provincia de Orense.

α_j/p	$p - \text{valores}$	
	Zaragoza	Orense
0.0042	0.0062	0.0011
0.0083	0.0149	0.0030
0.0125	0.0927	0.0100
0.0167	0.1083	0.0135
0.0208	0.1375	0.0367
0.0250	0.1544	0.0458
0.0292	0.1585	0.0906
0.0333	0.3419	0.1622
0.0375	0.4072	0.2370
0.0417	0.7324	0.2423
0.0458	0.9448	0.5088
0.0500	0.9756	0.5589
	Acepta	Rechaza

Tabla 4.21: Aplicación 2. Evaluación de si la nueva trayectoria pertenece al proceso original de temperaturas mínimas para Zaragoza y Orense.

De modo análogo, se ha realizado el test de bondad de ajuste sobre las

temperaturas máximas de dichas provincias, tras el que se obtienen los p-valores que se muestran en la tabla 4.22. Se puede observar, que se rechaza un comportamiento similar al modelo en algunas de las estaciones meteorológicas de clima más suave, Cádiz, Alcantarilla, Melilla, Málaga y Almería.

t	α_j/p	0.0042	0.0083	0.0125	0.0167	0.0208	0.0250	0.0292	0.0333	0.0375	0.0417	0.0458	0.0500
p -valores	Badajoz	0.0123	0.0129	0.0518	0.0566	0.0979	0.1335	0.1907	0.3647	0.6249	0.7388	0.8709	0.8962
	Córdoba	0.0171	0.0314	0.0615	0.0767	0.0777	0.1149	0.1180	0.1265	0.2054	0.2150	0.4094	0.5147
	Granada	0.0361	0.0557	0.1149	0.1180	0.2318	0.3748	0.3757	0.4906	0.5147	0.5796	0.7388	0.8962
	Sevilla	0.0051	0.0621	0.0622	0.0682	0.1265	0.1585	0.1907	0.3771	0.4906	0.4956	0.7388	0.8235
	Morón	0.0326	0.0557	0.0582	0.0841	0.1158	0.2318	0.3321	0.3757	0.5147	0.5187	0.8312	0.8674
	Rota	0.0283	0.0557	0.0622	0.0767	0.1180	0.1208	0.1529	0.1993	0.3964	0.5073	0.5187	0.7944
	Jerez	0.0043	0.0314	0.0320	0.1540	0.2134	0.3343	0.4113	0.5329	0.5509	0.6035	0.8075	0.9639
	Cádiz	0.0034	0.0063	0.0167	0.0557	0.2206	0.3389	0.3625	0.3683	0.3964	0.5329	0.8674	0.8709
	Alcantarilla	0.0010	0.0012	0.0070	0.0118	0.0122	0.0123	0.0518	0.0525	0.1046	0.3004	0.3375	0.5174
	Huelva	0.0085	0.0621	0.0974	0.1208	0.1661	0.2884	0.3964	0.5073	0.5187	0.6012	0.7691	0.8417
	Granada A	0.0051	0.0324	0.1149	0.1646	0.2134	0.3002	0.3113	0.3375	0.3757	0.3771	0.8195	0.8674
	Melilla	0.0036	0.0058	0.0077	0.0084	0.0139	0.0240	0.0299	0.0361	0.0442	0.1529	0.1661	0.3748
	Málaga	0.0067	0.0084	0.0086	0.0273	0.1020	0.1335	0.2123	0.2536	0.3113	0.3757	0.5509	0.8195
	Almeria	0.0018	0.0024	0.0058	0.0191	0.0442	0.1208	0.1661	0.1907	0.3113	0.3286	0.3995	0.4956
	Murcia	0.0146	0.0361	0.0566	0.0682	0.0726	0.3002	0.3375	0.3757	0.3964	0.6035	0.7388	0.7691

Tabla 4.22: Aplicación 2. P-valor para el test de bondad de ajuste sobre las temperaturas máximas.

En el caso de las temperaturas máximas y las provincias de Zaragoza y Orense, los resultados que proporciona el test de bondad de ajuste es que ninguna de las dos provincias contrastadas, sigue el modelo del proceso original, tal y como puede observarse en la tabla 4.23.

α_j/p	<i>p</i> – valores	
	Zaragoza	Orense
0.0042	0.0023	0.0001
0.0083	0.0171	0.0074
0.0125	0.0324	0.0075
0.0167	0.0582	0.0167
0.0208	0.1149	0.0301
0.0250	0.2206	0.0324
0.0292	0.2884	0.0557
0.0333	0.5073	0.0653
0.0375	0.5421	0.1150
0.0417	0.5716	0.5716
0.0458	0.7944	0.6035
0.0500	0.8312	0.8709
	Rechaza	Rechaza

Tabla 4.23: Aplicación 2. Evaluación de si la nueva trayectoria pertenece al proceso original de temperaturas máximas para Zaragoza y Orense.

4.4. Simulaciones del proceso auto-excitado

Pasamos en esta sección a abordar las simulaciones para el caso de los procesos auto-excitados, hay que recordar que en este caso existe cierta dependencia del pasado del proceso como ya se expuso en el Capítulo 3.

4.4.1. Simulación A-E caso 1

En la primera simulación se expone el caso de un proceso auto-excitado con intensidad una distribución tipo fase con representación $(\pi, T_{\omega_{i-1}^k})$ donde $\pi = (1, 0, 0)$ y

$$T = \begin{pmatrix} -10/\omega_{i-1}^k & 10/\omega_{i-1}^k & 0 \\ 0 & -10/\omega_{i-1}^k & 10/\omega_{i-1}^k \\ 0 & 0 & -10/\omega_{i-1}^k \end{pmatrix}$$

y donde los tiempos entre ocurrencias se distribuyen según una distribución exponencial

$$\omega_i^k - \omega_{i-1}^k \rightsquigarrow \exp(\lambda(\omega_{i-1}^k))$$

En este proceso, la intensidad $\lambda(t, x(t))$ depende de las ocurrencias antes de un tiempo t , $\omega_i^k < t$.

En primer lugar se han simulado los tiempos de ocurrencia ω_i^k a partir de $K = 100$ trayectorias, para continuar contando el número de ocurrencias del proceso $N(t)$, para ello el tiempo inicial ha sido establecido en $N(t_0) = 0$. Estos pasos previos, son necesarios para poder realizar la estimación de la intensidad condicionada al recuento de un cierto n , en un intervalo $[\text{mín}_k w_{n+1}^k, \text{máx}_k w_{n+1}^k]$.

La figura 4.18 muestra las simulaciones obtenidas de los tres primeros tiempos de ocurrencias, en las $K = 100$ trayectorias.

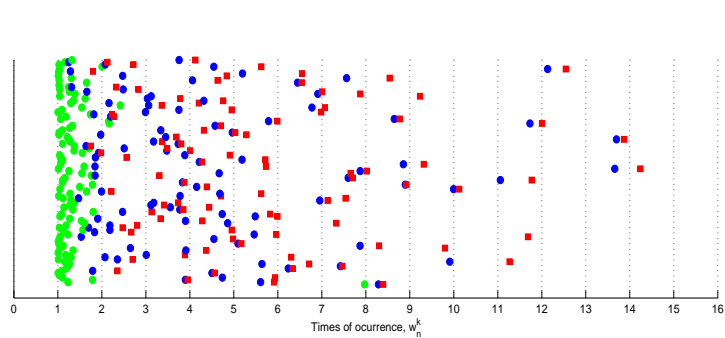


Figura 4.18: Simulación A-E 1. Tiempos de ocurrencias de las tres primeras ocurrencias.

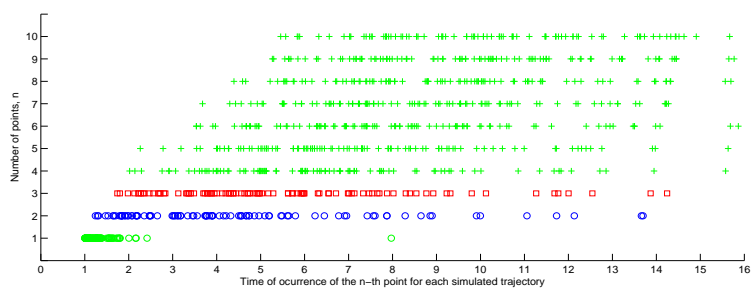


Figura 4.19: Simulación A-E 1. Estimación intensidad condicionada al recuento.

En el caso $n=0$,

$$\hat{\lambda}_0(t_j) = \frac{\sum_{k=1}^K \frac{N^k(t_j) - N^k(s)}{t_j - s} I \{N^k(t_j^-) = 0\}}{\sum_{k=1}^K I \{N^k(t_j^-) = 0\}}$$

el intervalo donde la función es distinta de 0, para la estimación de $\hat{\lambda}_0(t_j)$ resulta $t_j \in [\min_k w_1^k, \max_k w_1^k] \equiv [1.0042, 7.9721] \equiv [1.25, 7.75]$ y $s = t_{j-1}$. En la figura 4.20 se puede observar la representación del estimador puntual $\hat{\lambda}_0(t_j)$ así como del polinomio de interpolación cúbica. Análogamente, la estimación

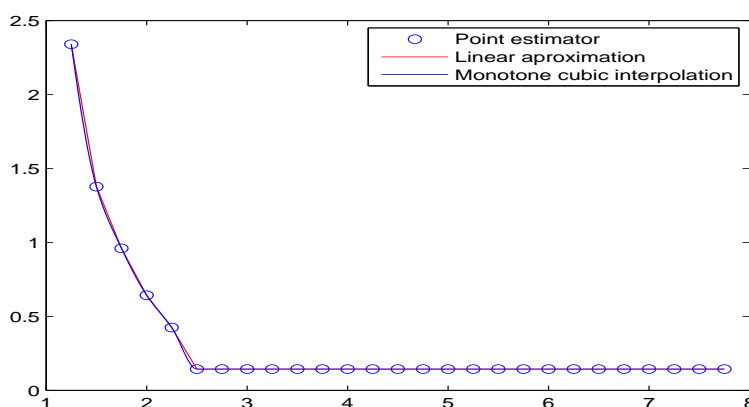


Figura 4.20: Simulación A-E 1. Estimación intensidad condicionada al recuento $n=0$.

para el caso $n=9$, es decir de $\hat{\lambda}_9(t_j)$, sería la siguiente

$$\hat{\lambda}_9(t_j) = \frac{\sum_{k=1}^K \frac{N^k(t_j) - N^k(s)}{t_j - s} I \{N^k(t_j^-) = 9\}}{\sum_{k=1}^K I \{N^k(t_j^-) = 9\}}$$

en el intervalo resultante $t_j \in [\min_k w_{10}^k, \max_k w_{10}^k] = [5.4551, 15.6663] \equiv [5.5, 14.75]$ y $s = t_{j-1}$. La figura 4.21 muestra la estimación puntual y la interpolación de $\hat{\lambda}_9(t_j)$

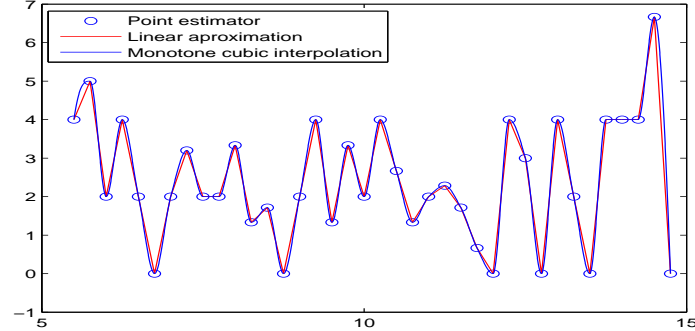


Figura 4.21: Simulación A-E 1. Estimación intensidad condicionada al recuento $n=9$.

En la tabla 4.24 se exponen los resultados de la estimación para la media del proceso, y de la media de la estimación de la intensidad, siguiendo las instrucciones de las ecuaciones (3.22) y (3.24)

La tabla 4.25 presenta la estimación de la función masa de probabilidad según las ecuaciones de la Sección 3.4.1.

4.4.2. Simulación A-E caso 2

En el segundo caso de simulación para los procesos auto-excitados, se considera un proceso con intensidad una distribución tipo fase con representación $(\pi, T_{\omega_{i-1}^k}) = E(3, 20/\omega_{i-1}^k)$, con $\pi = (1, 0, 0)$ y

$$T = \begin{pmatrix} -20/\omega_{i-1}^k & 20/\omega_{i-1}^k & 0 \\ 0 & -20/\omega_{i-1}^k & 20/\omega_{i-1}^k \\ 0 & 0 & -20/\omega_{i-1}^k \end{pmatrix}$$

los tiempos entre ocurrencias tienen distribución exponencial $\omega_i^k - \omega_{i-1}^k \rightsquigarrow \exp(\lambda(\omega_{i-1}^k))$ y la media de la intensidad es $E[\lambda(\omega_i^k)] = \frac{3\omega_{i-1}^k}{20}$.

En el desarrollo de este ejemplo se han simulado $K = 50$ trayectorias muestrales del proceso, representados en la figura 4.22

t	J	s	$E[N(t)]$	$E[\hat{\lambda}(t)]$
2.00	0	3	0.928	0.185
2.50	0	5	0.0023	0.0200
3.00	0	7	0.0000	0.0000
3.50	0	9	0.0000	0.0000
4.00	0	11	0.0000	0.0000
4.50	0	13	0.0000	0.0000
5.00	0	15	0.0000	0.0000
5.50	0	17	0.0000	0.0000
6.00	0	19	0.0000	0.0000
6.50	2	21	1.4954	0.4825
6.75	3	22	1.9859	0.02773
7.00	4	23	2.6234	0.2504
7.25	5	24	3.1028	0.2869
7.50	6	25	3.9105	0.2968
7.75	6	26	3.9709	0.1863
8.00	7	27	4.8397	0.2036
8.25	8	27	5.5252	0.0947
8.50	9	27	6.0503	0.0632
9.00	11	27	6.9752	0.0185
9.50	13	27	7.6000	0.0000
10.00	15	27	7.9203	0.0000

Tabla 4.24: Simulación A-E 1. Estimación de la media del proceso y de la media de la estimación de la intensidad.

$P[N(t) = n]$	t									
	1	2	3	4	5	6	7	8	9	10
n=0	1.000	0.186	0.142	0.123	0.106	0.092	0.080	0.0000	0.0000	0.0000
n=1	0.0000	0.188	0.217	0.236	0.253	0.267	0.279	0.0000	0.0000	0.0000
n=2	0.0000	0.027	0.079	0.072	0.067	0.065	0.065	0.064	0.064	0.064
n=3	0.0000	0.0000	0.045	0.071	0.076	0.089	0.077	0.057	0.081	0.081
n=4	0.0000	0.0000	0.011	0.027	0.062	0.043	0.056	0.096	0.105	0.076
n=5	0.0000	0.0000	0.0000	0.0000	0.028	0.044	0.057	0.075	0.070	0.083
n=6	0.0000	0.0000	0.0000	0.0000	0.021	0.048	0.046	0.045	0.059	0.063
n=7	0.0000	0.0000	0.0000	0.0000	0.0000	0.011	0.025	0.030	0.036	0.031
n=8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.009	0.025	0.028	0.045
n=9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.003	0.018	0.021	0.032
n=10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.003	0.015	0.018

Tabla 4.25: Simulación A-E 1. Estimación de la función masa de probabilidad.

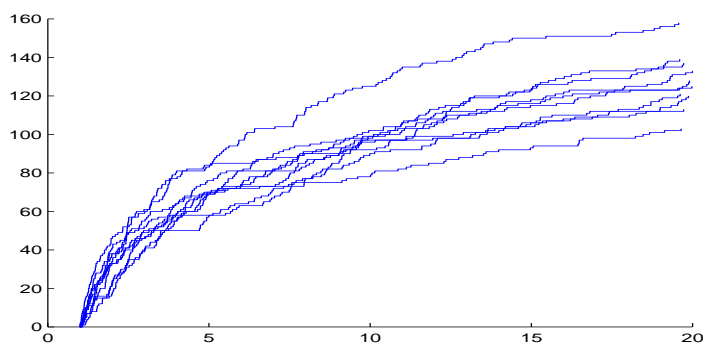


Figura 4.22: Simulación A-E 2. Trayectorias muestrales de proceso auto-excitado con intensidad tipo fase.

Siguiendo el mismo esquema de simulación del caso anterior, se obtiene que la simulación en las $K = 50$ trayectorias de las tres primeras ocurrencias del proceso siguen el esquema que se observa en la figura 4.23

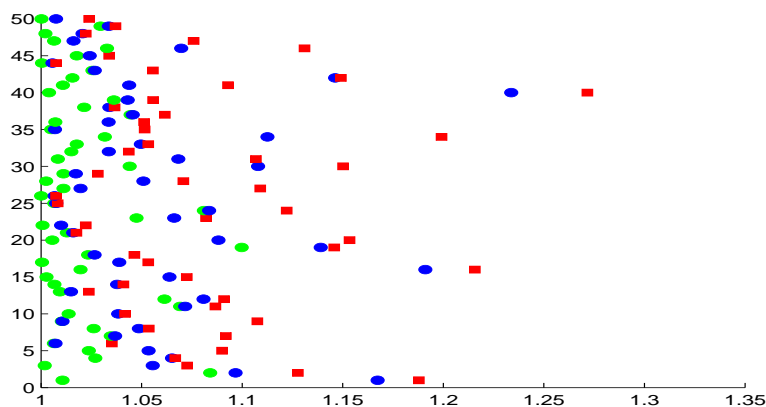


Figura 4.23: Simulación A-E 2. Tiempos de ocurrencias de las tres primeras ocurrencias.

Centrándonos en el caso $n = 2$, el estimador de la intensidad condicionada al recuento $\hat{\lambda}_n(t_j)$, puede verse en la figura 4.24.

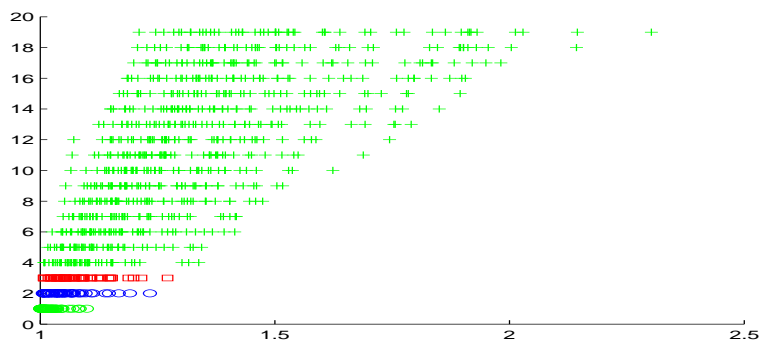


Figura 4.24: Simulación A-E 2. Estimación intensidad condicionada al recuento.

El intervalo donde la función es distinta de 0 para la estimación de $\hat{\lambda}_0(t_j)$ en el caso $n = 0$ resulta $t_j \in [\text{mín}_k w_1^k, \text{máx}_k w_1^k] \equiv [1.0057, 1.2338] \equiv [1.06, 1.23]$ y $s = t_{j-1}$. En la figura 4.25 se puede observar la representación

del estimador puntual así como del polinomio de interpolación cúbica.

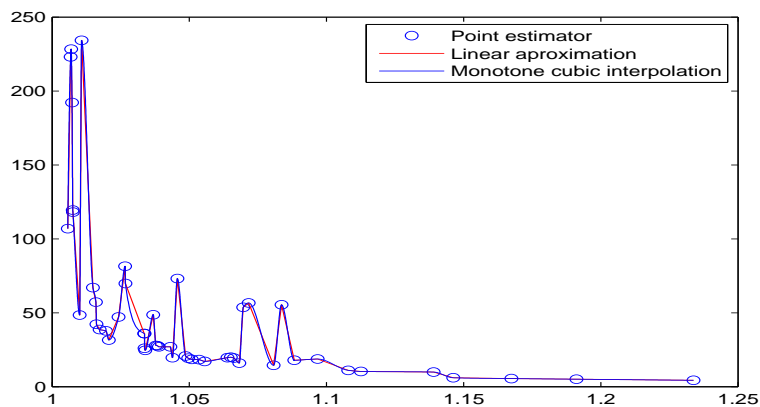


Figura 4.25: Simulación A-E 2. Estimación intensidad condicionada al recuento $n=1$.

Tras realizar la estimación de $\hat{\lambda}_9(t_j)$, el intervalo resultante $t_j \in [\text{mín}_k w_{10}^k, \text{máx}_k w_{10}^k] = [1.068, 1.6885] \equiv [1.07, 1.69]$, $s = t_{j-1}$. La figura 4.26 muestra dicha estimación.

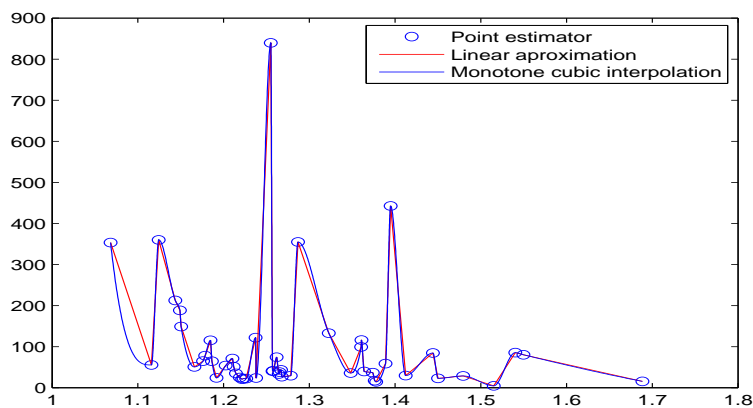


Figura 4.26: Simulación A-E 2. Estimación intensidad condicionada al recuento $n=9$.

Apéndice A

Conclusiones

La presente memoria proponía como objetivo genérico “el estudio de procesos de recuento doblemente estocásticos partiendo del de Cox, tanto en su componente descriptiva (estudio de sus estadísticos) como en cuanto a su inferencia (estimación y contraste de hipótesis) desde la perspectiva clásica de Snyder y Miller, 1991 [117]”. A lo largo del texto se ha ido trabajando y respondiendo a ese objetivo genérico y a cada uno de los específicos detallados en Objetivos. Con ánimo de no ser repetitivos, se recogen ahora de forma sistemática los resultados y conclusiones indicando dónde encontrar cada uno de ellos y atendiendo a dos bloques, el primero, referido a los PC y PC compuesto, y el segundo, dedicado a los procesos auto-excitados. Además se señalan las referencias que avalan el trabajo y en las que ha participado la autora de esta memoria.

A.1. Proceso de Cox y Cox compuesto

Estos procesos se definen en los Capítulos 1 y 2, respectivamente. El PC compuesto estaba falto de un mayor desarrollo teórico por lo que se aportan

expresiones cerradas para sus estadísticos de recuento en la Sección 2.3, y sus estadísticos de tiempo en la Sección 2.4. La importancia de estos resultados, va más allá de la obviedad de su mera aportación, pues abre la posibilidad de desarrollar inferencia para el proceso.

Siguiendo esta dirección, en los epígrafes 2.5.1 y 2.5.2, se extiende la estimación presentada para el PC a esta generalización, el PC compuesto. Esto implica que se ha usado la misma filosofía de estimar sin haber hecho suposiciones previas de la estructura del proceso, lo que es un punto fuerte de la metodología propuesta.

Tanto para el PC como para el PC compuesto, también se había planteado como objetivo, construir un contraste de bondad de ajuste para decidir si, habiendo observado una trayectoria de un proceso de recuento, se podía asumir perteneciente a un PC o PC compuesto ya conocido u observado previamente. Los epígrafes 1.4.3 y 2.5.3 presentan los respectivos contrastes para resolver ese problema.

Las expresiones de los estadísticos y sus correspondientes estimaciones se ilustran y evalúan por medio de varias simulaciones en la Sección 4.2. En el epígrafe 4.2.3, se muestra también cómo aplicar el contraste de bondad de ajuste y su evaluación. En todas las simulaciones se presenta la comparación entre las estimaciones, basadas en técnicas conocidas por su capacidad y precisión (análisis en componentes principales), y los valores muestrales, quedando patente el buen comportamiento de todos y cada uno de los estimadores. De la misma manera ocurre con el test de bondad de ajuste; la evaluación de su precisión deja clara su capacidad de discernir entre nuevos datos observados y su correspondencia o no con el proceso patrón.

Habiendo concluído que los estimadores y el test son buenas herramientas, todos esos resultados teóricos se aplican a datos reales en la Sección 4.3. Como

se vió entonces, las estimaciones muestran su buen comportamiento y el test de bondad de ajuste arroja resultados perfectamente razonables.

Los resultados y conclusiones comentadas han sido presentados en congresos internacionales (Ruiz-Fuentes et al., 2011 [103], Ruiz-Fuentes et al., 2013 [104], Ruiz-Fuentes et al., 2014 [105], Montes-Gijón et al., 2016 [87], Ruiz-Fuentes et al., 2016 [106], Bouzas et al., 2017 [32], Ruiz-Fuentes et al., 2017a [107], Ruiz-Fuentes et al., 2017 [108]) y publicados en el artículo Bouzas et al., 2019 [33].

A.1.1. Proceso tipo fase

Es necesario hacer mención especial a otra aportación de esta memoria, el proceso tipo fase y su uso como proceso de intensidad de un proceso de recuento.

En otro intento de generalizar los PC, la búsqueda de procesos intensidad que recojan la mayor variedad posible de situaciones es fundamental. La clase de distribuciones tipo fase es densa en las distribuciones de probabilidad no-negativas continuas, por tanto resulta un tipo de distribuciones idóneas para modelizar una variable no-negativa. En la Definición 1.1.4 (pág. 11) se define el que se ha denominado proceso tipo fase, como un proceso basado en una distribución tipo fase cuyos parámetros son a su vez aleatorios. Con un razonamiento paralelo al dado para variables no-negativas, este nuevo proceso será un modelo ideal para un proceso intensidad, ya no solo de un PC sino de un proceso de recuento cualquiera.

En conclusión, se ha definido un proceso no-negativo muy flexible y general, características que hereda cualquier proceso de recuento que lo tome como proceso intensidad.

Además, se resolvió el reto que se presentaba, la simulación de este nuevo

proceso y su inclusión como intensidad de un proceso de recuento, lo que se ha visto que no siempre es factible y, por tanto, ha abierto una nueva línea de estudio. Se ha simulado un ejemplo de proceso tipo fase en la pág. 11, que más adelante, en el epígrafe 4.2.3, se utiliza como proceso intensidad de un PC compuesto. También se simulan otros más para usar como intensidades de procesos auto-excitados en los epígrafes 4.4.1-??.

El proceso tipo fase fue definido formalmente por primera vez en Bouzas et al., 2017 [32] y se ha publicado en el artículo Bouzas et al., 2019 [33].

A.2. Proceso auto-excitado

El proceso de recuento auto-excitado se define en el Capítulo 3, como PC que incluyen dependencia del pasado del proceso, ya sea del recuento, como de los tiempos de ocurrencia.

Si era difícil encontrar expresiones explícitas para los PC compuesto, más aún para estos otros. En la Sección 3.2 se proporcionan expresiones explícitas de los estadísticos de recuento. Dado que todos ellos dependen de la llamada intensidad condicionada al recuento, resulta obvio que una estimación de la misma, proporcionaría la posibilidad de estimar los estadísticos. En la Sección 3.4 se construye un estimador puntual para la intensidad condicionada al recuento y, a partir de ella, las estimaciones de la función masa de probabilidad, de la media, y de la intensidad media.

Todos los resultados teóricos se ilustran por medio de aplicaciones de los mismos a simulaciones en la Sección 4.4. Estas simulaciones presentaron varias dificultades, la inclusión de un proceso tipo fase como intensidad con algún tipo de dependencia del pasado, la aplicación de la fórmula de estimación de la intensidad condicionada y las integrales en las estimaciones de los

estadísticos resueltas mediante cálculo numérico y el hecho de que aparecieran encadenadas en la estimación de la función masa de probabilidad.

En conclusión, se aportan resultados teóricos interesantes, se construyen simulaciones y se da solución a problemas prácticos de cálculo que fueron presentados en Ruiz-Fuentes et al., 2014 [105].

Apéndice B

Resumen

Un proceso puntual es el resultado de modelizar las distintas ocurrencias de cualquier fenómeno en el cual un suceso se repite en el tiempo o el espacio. Ligado a todo proceso puntual, se encuentra el proceso de recuento, que tal y como su propio nombre indica, se centra en contar el número de puntos, sucesos u ocurrencias en intervalos del tiempo o el espacio. Las distribuciones de probabilidad discretas, juegan un papel primordial en el desarrollo de este tipo de procesos, siendo la distribución de Poisson, hallada por Siméon Denis Poisson, el origen del punto de vista usado en esta tesis sobre los procesos de recuento.

El interés en el estudio de estos procesos ha ido creciendo a lo largo del tiempo, siendo quizá las tres áreas en las que su aplicación es más significativa, la teoría de colas, la teoría actuarial y la teoría de fiabilidad. Tras la Segunda Guerra Mundial se produjo un gran desarrollo tanto en la teoría como en las aplicaciones de los procesos estocásticos, surgiendo de modo paralelo el desarrollo matemático para los procesos puntuales en un espacio de estados general (ver Moyal, 1962 [88]) y el artículo de Cox, 1955 [40] en el que por vez primera se cita el proceso de Poisson doblemente estocástico, también

conocido como proceso de Cox. Este nuevo proceso es una generalización de los anteriores, cuya particularidad es que su intensidad es estocástica, lo cual le confiere una mayor flexibilidad. A partir de esta publicación se incrementa el interés por los procesos puntuales tanto en comunicaciones como en medicina, focalizándose en la modelización y predicción de los procesos.

En la última parte del siglo XX se ha consolidado en diversos campos el uso de procesos doblemente estocásticos, si bien el ejemplo más utilizado para ilustrar la aplicación de éstos a datos reales, suele ser el recuento de terremotos (ver Ogata, 1998 [92], Gospodinov y Pavlova, 2001 [58]). Aunque han sido diversos los autores que se han dedicado al estudio de los procesos de Cox, es de destacar el estudio realizado por Snyder y Miller, 1991 [117], que es quizá el texto más completo que muestra la evolución desde el proceso de Poisson hasta el proceso de Cox.

Cierto es que la estimación del proceso de Cox ha sido estudiada por distintos autores, aunque siempre asumiendo cierta estructura en la intensidad del proceso. Esta memoria centra su interés en el desarrollo de Bouzas, 2011 [28], en el cual utiliza la técnica de análisis funcional en componentes principales para las estimaciones y predicciones del proceso, siendo esta técnica bien conocida, de gran potencial y precisión. Bouzas y Ruiz-Fuentes, 2015 [31] realiza un repaso de la inferencia hecha usando dicha técnica hasta ese momento tanto en el proceso de Cox, como en el proceso de Cox compuesto. Esta memoria se centra en los procesos doblemente estocásticos, en su estudio y la necesidad de su continua generalización con el fin de evitar cualquier restricción sobre el modelo y modelizar fielmente casos reales.

B.1. Capítulo 1: Proceso de Cox

El primer capítulo de la memoria aborda el estudio del proceso de Poisson cuya intensidad se ha aleatorizado, pasando a ser un proceso de Poisson doblemente estocástico, conocido también como proceso de Cox (PC), dado que el primer estudio detallado de este tipo de procesos se realizó en Cox, 1955 [40]. Más tarde, Snyder y Miller, en su versión de 1991 [117], fue quien llevó a cabo el estudio más amplio del proceso de Poisson y sus generalizaciones, tales como el proceso de Poisson compuesto o el proceso de Poisson filtrado, y, ya en su último capítulo, presenta el PC y alguna de sus generalizaciones de modo breve.

Al margen de la definición del PC teniendo de base el proceso de Poisson (ver Snyder y Miller, 1991 [117]), que es bastante intuitiva, en esta memoria, se hace uso también de la definición dada por Serfozo, 1972 [112], basada en la media del proceso. El PC suele tener, tal y como se muestra en el Capítulo 1, expresiones para sus estadísticos, de gran complejidad. Todos ellos dependen del proceso intensidad o del proceso media, por lo que la profundización en el estudio de uno u otro es fundamental tanto para el desarrollo como para la clarificación de los estadísticos, y más adelante para su inferencia.

Es habitual encontrar en la literatura modelos sujetos a gran cantidad de restricciones; en la memoria, excepto para ejemplos o simulaciones, no se han utilizado modelos con intensidad o media concretas. No obstante, uno de los objetivos de este capítulo, es definir un modelo que permita tener una intensidad general, flexible y sin restricciones, por lo que se ha hecho uso de las distribuciones tipo fase. Estas distribuciones se caracterizan por ser no-negativas y de soporte finito, por lo que son ideales para modelizar la intensidad de cualquier proceso estocástico. No hay que olvidar que la

intensidad del PC es a su vez un proceso estocástico, por lo que se ha dado un paso, aleatorizando los parámetros de la distribución tipo fase, obteniéndose el proceso tipo fase que se define de la siguiente manera.

Un **proceso tipo fase** es un proceso estocástico con representación $(\boldsymbol{\pi}, \mathbf{T})$, donde

1. $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_p)_{1 \times p}$ y $\mathbf{T} = (\mathbf{T}_{ij})_{p \times p}$
2. siendo los parámetros $\boldsymbol{\pi}$ y/o \mathbf{T} variables aleatorias.
3. para cada realización de la variable aleatoria, su distribución es una distribución tipo fase.

Con este proceso se consigue obtener un proceso intensidad general, flexible, que no necesita restricciones, válido para modelizar la intensidad de cualquier proceso estocástico de todo campo de estudio, por lo que es una herramienta muy útil.

A lo largo del Capítulo 1 se hace una revisión tanto de los estadísticos de recuento como de los estadísticos de tiempo. En el apartado de inferencia, se resumen los Análisis de Componentes Principales Funcional (ACPF) *ad hoc* para la estimación del proceso media y del proceso intensidad, para después predecir mediante los modelos de Predicción en Componentes Principales (PCP). Estas herramientas permiten dar la siguiente expresión para estimar la media del PC:

$$\tilde{\Lambda}^{q_2}(s) = \mu_{\Lambda}^2(s) + \sum_{j=1}^{q_2} \left(\sum_{i=1}^{p_j} b_i^j \xi_i \right) g_j(s); \quad s \in (T_1, T_2)$$

y para la intensidad del PC:

$$\lambda^q(t) = p'^q(t) = \mu_{p'}(t) + \sum_{j=1}^q \zeta_j f_j(t)$$

a partir de la cual pueden obtenerse estimaciones de los estadísticos del PC, puesto que todos ellos se expresan en función de alguna de ellas, o de ambas.

El otro objetivo del capítulo se centra en saber si ciertos datos provienen de un proceso estocástico concreto, es decir, en la realización de un contraste de bondad de ajuste. Realizar dicho tipo de contraste sobre datos que provienen de un proceso estocástico tiene la dificultad añadida de que cada trayectoria muestral es una función. Se aborda el problema desde dos escenarios distintos: el contraste de bondad de ajuste dado un PC conocido, y el contraste de bondad de ajuste sobre datos observados de un PC del cual se desconoce su estructura.

B.1.1. Contraste de bondad de ajuste dado un PC

En el caso en el que la estructura estocástica del PC es conocida, el estimador puntual (insesgado y consistente) de la intensidad viene dado por:

$$\hat{\lambda}_l(t_j) = \frac{1}{m} \sum_{i=1}^m \frac{M_i(t_j) - M_i(t_{j-1})}{t_j - t_{j-1}}, \quad j = 1, \dots, p; \quad l = 1, \dots, s$$

El intervalo de confianza para la media si la intensidad es conocida y gaussiana o si $s \geq 30$, el intervalo de confianza para $\mu(t_j)$

$$I_j = \left[\mu_{exp}(t_j) \pm z_{1-\alpha/2} \sqrt{\frac{1}{s} \sigma^2(t_j)} \right], \quad j = 1, \dots, p$$

donde $z_{1-\alpha/2}$ es el percentil $(1 - \alpha/2)$ de una distribución Normal estándar.

El contraste para un $\mu_{exp}(t_j)$ fijo es:

$$\begin{cases} H_0 : E[\lambda(t_j)] = \mu(t_j) \\ H_1 : E[\lambda(t_j)] \neq \mu(t_j) \end{cases}$$

cuyo p-valor es:

$$p\text{-valor}(t_j) = 2P \left[Z > \left| \frac{\mu_{exp}(t_j) - \mu(t_j)}{\sqrt{\frac{1}{s} \sigma^2(t_j)}} \right| \right]$$

Por tanto H_0 no se rechaza si $p - valor(t_j) > \alpha$. Para considerar si una trayectoria sigue un cierto PC, no serviría solo con estudiar puntos t_j de forma independiente. Por este motivo, se propone un test múltiple con t_j , $j = 1, \dots, p$ y el uso de la inferencia simultánea con el criterio de Benjamini y Hochber, 1995 [13], el test resulta ser el siguiente

$$\begin{cases} H_0^{(j)} : E[\lambda(t_j)] = \mu(t_j) \\ H_1^{(j)} : E[\lambda(t_j)] \neq \mu(t_j) \end{cases}, \quad j = 1, \dots, p$$

y los p-valores correspondientes $p - valor^{(1)} \leq \dots \leq p - valor^{(p)}$, H_0 se acepta si y solo si $p - valor^{(1)} > \alpha/p$, $p - valor^{(2)} > 2\alpha/p, \dots, p - valor^{(j)} > j\alpha/p, \dots, p - valor^{(p)} > \alpha$.

B.1.2. Contraste de bondad de ajuste dado los datos observados de un PC

Es habitual que la intensidad del PC no sea conocida de modo paramétrico. Por ello se aborda el caso en el que lo único que se conoce del PC son trayectorias muestrales observadas. En este caso el intervalo de confianza para la media es:

$$\left[\mu_{exp}(t_j) \pm t_{1-\alpha/2, s-1} \sqrt{\frac{1}{s} \hat{\sigma}_j^2} \right]$$

donde $t_{1-\alpha/2, s-1}$ es el percentil $(1 - \alpha/2)$ de una distribución t de student. Siendo en este caso el contraste:

$$\begin{cases} H_0 : E[\lambda(t_j)] = \hat{\mu}(t_j) \\ H_1 : E[\lambda(t_j)] \neq \hat{\mu}(t_j) \end{cases}$$

y el p-valor

$$p - valor(t_j) = 2P \left[t > \left| \frac{\mu_{exp}(t_j) - \hat{\mu}(t_j)}{\sqrt{\frac{1}{s} \hat{\sigma}_j^2}} \right| \right]$$

donde t es una variable distribuida según una t de student con $s - 1$ grados de libertad.

Análogamente al caso anterior, se propone un contraste múltiple con el mismo criterio de aceptación.

$$\begin{cases} H_0^{(j)} : E[\lambda(t_j)] = \hat{\mu}(t_j) \\ H_1^{(j)} : E[\lambda(t_j)] \neq \hat{\mu}(t_j) \end{cases}, \quad j = 1, \dots, p$$

B.2. Capítulo 2: Proceso de Cox compuesto

El Capítulo 2, aborda una generalización del PC, el PC compuesto, un proceso puntual cuyas ocurrencias tienen asociada una variable aleatoria. En el caso particular que nos ocupa, el proceso es un PC y las marcas son variables aleatorias independientes e idénticamente distribuidas, e independientes a su vez del PC. Este proceso es una ampliación del proceso de Poisson compuesto estudiado por Snyder y Miller, 1991 [117]. El PC compuesto puede modelizar un gran número de sucesos distintos, aunque es especialmente interesante la interpretación en la que se estudia el número de puntos ocurridos en cada instante. El hecho de permitir la ocurrencia simultánea de puntos, relaja la propiedad de regularidad, lo cual hace que sea un modelo más flexible para su uso en diversos campos de estudio.

El desarrollo del PC compuesto se basa en los teoremas de representación, expuestos en los teoremas 2.2.1 y 2.2.2, a partir de los cuales se definen los estadísticos de recuento del PC compuesto en la Sección 2.3, y se aportan expresiones cerradas de sus estadísticos de tiempo, tales como,

- **Función de supervivencia**

$$P^B[W_n > T] = P[N(t, B) < n]$$

- **Función de densidad condicionada**

$$\begin{aligned}
 & p_{w_n/w_{n-1}}^B(W_n/W_{n-1}) \\
 &= E_x \left[\lambda(W_n) \int_B P_u(dU) \exp \left[- \int_{W_{n-1}}^{W_n} \lambda(\sigma) P_u(dU) d\sigma \right] \right] \\
 &= E_x \left[\lambda(W_n) \int_B P_u(dU) \exp \left[-(\Lambda(W_n) - \Lambda(W_{n-1})) \int_B P_u(dU) \right] \right]
 \end{aligned}$$

- **Probabilidad de tener una nueva ocurrencia en un intervalo de tiempo**

$$\begin{aligned}
 & P_{w_n/w_{n-1}}^B(W_n < T/W_{n-1}) \\
 &= E_x \left[1 - \exp \left[-(\Lambda(T) - \Lambda(W_{n-1})) \int_B P_u(dU) \right] \right]
 \end{aligned}$$

- **Función de supervivencia condicionada**

$$\begin{aligned}
 & P_{w_n/w_{n-1}}^B(W_n > T/W_{n-1}) = \\
 & E_x \left[\exp \left[-(\Lambda(T) - \Lambda(W_{n-1})) \int_B P_u(dU) \right] \right]
 \end{aligned}$$

- **Densidad de ocurrencia adelantada**

$$\begin{aligned}
 & p_{t_n/w_{n-1}}^B(T/W_{n-1}) = p_{w_n/w_{n-1}}^B(W_n + T/W_{n-1}) = \\
 & E_x \left[\lambda(W_n + T) \int_B P_u(dU) \exp \left[-(\Lambda(W_n + T) - \Lambda(W_{n-1})) \int_B P_u(dU) \right] \right]
 \end{aligned}$$

En la Sección 2.5, dedicada a la inferencia sobre el PC compuesto, se proporcionan estimadores para los estadísticos de recuento introducidos en la estimación del proceso medio. Sin embargo, la estimación de la probabilidad de que haya una nueva ocurrencia antes de un cierto instante, resulta intratable. La proposición 2.5.1 resuelve el problema bajo ciertas restricciones.

Proposición

Dada la estimación del proceso media de $N(t)$ de la ecuación (1.29), se asume que las componentes principales ξ_i utilizadas para la estimación de las componentes del futuro η_j , $j = 1, \dots, q_2$ son independientes y se distribuyen según una distribución Normal (o $k > 30$) y sus distribuciones son (o pueden estimarse como) $\mathcal{N}(0, \sigma_i^2)$ donde σ_i^2 son las varianzas estimadas previamente en el correspondiente ACPF.

Para un tiempo fijo $s \in (W_{n-1} = T_1, T_2)$, la probabilidad de haber una nueva ocurrencia con marca en B en el intervalo $(W_{n-1} = T_1, s)$ es

$$P_{w/w_{n-1}}^B(W_n < s/W_{n-1}) \simeq 1 - \exp \left[\left(\Lambda^{q_1}(W_{n-1}) - \mu_{\Lambda}^2(s) \right) \int_B P_u(dU) \right] \cdot \exp \left[\frac{\left(\int_B P_u(dU) \right)^2}{2} \sum_{j=1}^{q_2} \left(\sum_{i=1}^p (\sigma_i b_i^j)^2 \right) (g_j(s))^2 \right]$$

donde p es el mayor p_j , $\forall j$ ■

B.2.1. Contraste de bondad de ajuste de un PC compuesto con marcas específicas

De modo análogo al caso del PC, se estudia el contraste de bondad de ajuste de un PC compuesto con marcas específicas. La generalización es clara puesto que el proceso de recuento sigue siendo un PC. Por lo que tomando nuevas trayectorias de un nuevo proceso puntual $M(t)$, se puede comprobar si el proceso formado por puntos con marcas en B , $\{M(t, B); t \geq t_0\}$, es o no un PC con intensidad $\left\{ \lambda^B(t) = \lambda(t) \int_B P_u(dU); t \geq t_0 \right\}$, según los teoremas de representación. De nuevo se distinguen dos situaciones distintas a la hora de realizar el contraste, el caso en el que se conoce el proceso, y aquel en el cual se realiza a partir de datos observados del mismo, ya que la estructura

estocástica no es conocida. En ambos el criterio de aceptación o rechazo es el criterio BH visto en el capítulo anterior.

Contraste de bondad de ajuste dado un PC compuesto

Puesto que la intensidad del proceso es conocida, el test a realizar sería

$$\begin{cases} H_0^{(j)} : E[\lambda^B(t_j)] = \mu(t_j) \int_B P_u(dU) \\ H_1^{(j)} : E[\lambda^B(t_j)] \neq \mu(t_j) \int_B P_u(dU) \end{cases}, \quad j = 1, \dots, p$$

donde $\sigma_{\lambda^B}^2(t) = \left[\int_B P_u(dU) \right]^2 \sigma^2(t)$. El p-valor para el instante t_j es:

$$p - \text{valor}(t_j) = 2P \left[Z > \left| \frac{\mu_{exp}(t_j) - \int_B P_u(dU) \mu(t_j)}{\int_B P_u(dU) \sqrt{\frac{1}{s} \sigma^2(t_j)}} \right| \right]$$

Contraste de bondad de ajuste dado los datos observados de un PC compuesto

En el caso en el que se desconoce la estructura estocástica del PC compuesto, se realiza el test a partir de datos observados del mismo, siendo el contraste a realizar:

$$\begin{cases} H_0^{(j)} : E[\lambda^B(t_j)] = \hat{\mu}(t_j) \int_B P_u(dU) \\ H_1^{(j)} : E[\lambda^B(t_j)] \neq \hat{\mu}(t_j) \int_B P_u(dU) \end{cases}, \quad j = 1, \dots, p$$

es necesario uso del estimador puntual de la intensidad (1.34) para estimar la intensidad entre los puntos con marcas en B . Nótese que ahora $\hat{\sigma}_{\lambda^B}^2(t) = \left[\int_B P_u(dU) \right]^2 \hat{\sigma}^2(t)$. El p-valor del contraste en t_j es

$$p - \text{valor}(t_j) = 2P \left[t > \left| \frac{\mu_{exp}(t_j) - \int_B P_u(dU) \hat{\mu}(t_j)}{\int_B P_u(dU) \sqrt{\frac{1}{s} \hat{\sigma}^2(t_j)}} \right| \right]$$

B.3. Capítulo 3: Procesos auto-excitados

El Capítulo 3, da un paso más hacia adelante, buscando la mayor generalización posible en el campo de los procesos doblemente estocásticos. Relajando la propiedad de la falta de memoria del PC, surge el interés por los procesos auto-excitados, procesos que tienen cierta dependencia del pasado del proceso, como es el caso de la emisión de radiación en isótopos radiactivos de vida corta, en los cuales la radiación no es la misma al inicio del proceso que transcurrido un cierto tiempo.

El caso más general de proceso auto-excitado, es aquel en el que toda ocurrencia y todo tiempo de ocurrencia anterior, puede influir en el número y tiempo de ocurrencia de todo punto posterior. Sin embargo, los procesos más conocidos de este tipo, son aquellos en los que la dependencia es solo de alguna parte o fracción del proceso, como los últimos tiempos de ocurrencia, o el número total de puntos acumulados, ejemplo de ello son los procesos de renovación y los procesos de nacimiento de Markov.

Aunque son utilizados en muchos artículos, los ejemplos encontrados en la literatura, no son tan generales como cabría esperar. El estudio más detallado que se encuentra hasta ahora es el desarrollado por Snyder y Miller, 1991 [117].

En esta memoria, a parte de un estudio pormenorizado de los procesos auto-excitados, en el que a parte de su definición, se muestran sus estadísticos de recuento en el epígrafe 3.2, y los estadísticos de tiempo en el epígrafe 3.3. En la Sección 3.4 se propone una expresión para la estimación de la intensidad

condicionada al recuento

$$\widehat{\lambda}_n(t_j) = \frac{\sum_{k=1}^K \frac{N^k(t_j) - N^k(\omega_n^k)}{t_j - \omega_n^k} I \{N^k(t_j^-) = n\}}{\sum_{k=1}^K I \{N^k(t_j^-) = n\}}, \quad n \in \mathbb{N}$$

para $t_j < T$ y donde t_j^- es el tiempo al que tiende t_j por la izquierda, e I es la función indicadora.

Este estimador es clave para poder estimar algunos estadísticos de recuento del proceso auto-excitado, como la función masa de probabilidad vista en el epígrafe 3.4.1, la media mostrada en la ecuación (3.22) y la intensidad media expuesta en la ecuación (3.24).

B.4. Capítulo 4: Simulaciones y aplicaciones

El último capítulo pone en práctica el desarrollo teórico llevado a cabo a lo largo del documento. El motivo de realizar un capítulo especialmente para ello en lugar de desarrollarlo a lo largo de la memoria, es porque tanto en las simulaciones como en las aplicaciones se utilizan resultados obtenidos en distintos capítulos, por lo que se ha considerado que unificándolo se obtenía una presentación más clara y completa de los mismos.

Mediante el software MATLAB R2016b, se han construido las simulaciones tanto para el PC compuesto vistas en la Sección 4.2, como para el proceso auto-excitado en la Sección 4.4, así como para la aplicación de los resultados teóricos.

Con el fin de poner de manifiesto la utilidad de los PC y PC compuestos en la modelización de casos reales, los resultados teóricos también se han utilizado sobre datos reales en la Sección 4.3. En concreto un primer caso sobre puntos de cambio de tendencia en datos bursátiles expuesto en el epígrafe

4.3.1, y una segunda aplicación sobre datos extremos en meteorología desarrollado en el epígrafe 4.3.2, obteniéndose en ambos casos resultados muy satisfactorios, tal y como se muestra en los epígrafes dedicados a dichas aplicaciones, por lo que queda expuesta la gran utilidad del estudio y desarrollo teórico realizado a lo largo de la presente memoria.

Bibliografía

- [1] Aalen O, Borgan Ø, Gjessing HK. **Survival and event history analysis: A process point of view**. Statistics for Biology and Health, New York, 2008.
- [2] Abbé E. Über Blutkörper-Zählung. **Jenaische Zeitschrift für Naturwissenschaft** (in German). Jena, Germany: Verlag Von Gustav Fischer, 98-105, 1879.
- [3] Agencia Estatal de Meteorología (AEMET). **Atlas climático ibérico. Temperaturas del aire y precipitación (1971-2000)**. Ministerio de Medio Ambiente y Medio Rural y Marino, 2011.
- [4] Aguilera AM, Ocaña FA, Valderrama MJ. **An approximated principal component prediction model for continuous-time stochastic processes**. Applied Stochastic Models and Data Analysis, 13, 61-72, 1997.
- [5] Al Mutairi D, Chen Y, Singpurwalla ND. **An adaptative concatenated failure rate model for software reliability**. Journal of the American Statistical Association, 93, 443, 1150-1163, 1998.
- [6] Andersen PK, Borgan Ø, Gill RD, Keiding N. **Statistical models based on counting processes**. Springer-Verlag, New York, 1993.

- [7] Asmussen S. **Phase type representations in random walk and queuing problems**. The Annals of Probability, 30, 2, 772-789, 1992.
- [8] Asmussen S. **Applied probabilities and queues**. Springer-Verlang, New York, 2003.
- [9] Asha G, Nair UN. **Reliability properties of mean time to failure in age replacement models**. International Journal of Reliability, Quality and Safety Engineering 17, 15-26, 2010.
- [10] Barta P, Miller Ml, Qiu A. **A stochastic model for studying the laminar structure of cortex from MRI**. IEEE Trans Med Imaging 24, 728-742, 2005.
- [11] Bateman, H. **Note on the probability distribution of α -particles**. Philosophical Magazine, 20, 704-707, 1910.
- [12] Bauwens L, Hautsch N. **Modelling financial high frequency data using point processes**. Handbook of financial time series, Springer-Verlag, Berlin, Heidelberg, 953-979, 2009.
- [13] Benjamini Y, Hochberg Y. **Controlling the false discovery rate: a practical and powerful approach to multiple testing**. Journal of the Royal Statistical Society, Statistical Methodology Series B, 57, 289-300, 1995.
- [14] Bieniek M, Goroncy A. **Sharp lower bounds on expectations of gOS based on DGFR distributions**. Statistical Papers, Springer, 1-16, 2017.

-
- [15] Bladt M. **A review on Phase-type distributions and their use in risk theory**. ASTIN Bulletin: The Journal of the IAA, 35, 1, 145-161, 2005.
- [16] Borak S, Detlefse K, Härdle W. **FFT based option pricing**. Statistical Tools for Finance and Insurance, 183-200, 2005.
- [17] Boel R, Beneš V. **Recursive nonlinear estimation of a diffusion acting as the rate of an observed Poisson process**. IEEE Transactions on Information Theory, 26, 561-575, 1980.
- [18] Bouzas PR. **Predicción con modelos estocásticos de tipo poissoniano**. Tesis doctoral de la Universidad de Granada. 1999.
- [19] Bouzas PR, Aguilera del Pino AM, Valderrama MJ. **Forecasting a class of doubly stochastic Poisson processes**. Statistical Papers, 43, 507-523. 2002.
- [20] Bouzas PR, Valderrama MJ, Aguilera AM. **A theoretical note on the distribution of a filtered compound doubly stochastic Poisson process**. Applied Mathematical Modeling 28, 769-773, 2004.
- [21] Bouzas PR, Valderrama MJ, Aguilera AM. **On the characteristic functional of a doubly stochastic Poisson process: Application to a narrow-band process**. Applied Mathematical Modelling, 30, 1021-1032, 2006a.
- [22] Bouzas PR, Aguilera AM, Valderrama MJ, Ruiz-Fuentes N. **On the structure of the stochastic process of mortgages in Spain**. Computational Statistics, 21, 73-89, 2006b.

- [23] Bouzas PR, Valderrama MJ, Aguilera AM, Ruiz-Fuentes N. **Modelling the mean of a doubly stochastic Poisson process by functional data analysis**. Computational Statistics and Data Analysis, 50, 2655-2667, 2006c.
- [24] Bouzas PR, Ruiz-Fuentes N, Ocaña FM. **Functional approach to the random mean of a compound Cox process**. Computational Statistics, 22, 467-479, 2007.
- [25] Bouzas PR, Aguilera AM, Ruiz-Fuentes N. **Functional estimation of the random rate of a Cox process**. Methodology and Computing in Applied Probability, On line, DOI: 10.1007/s11009-010-9173-z, 2010a.
- [26] Bouzas PR, Ruiz-Fuentes N, Matilla A, Aguilera AM, Valderrama MJ. **A Cox model for radioactive counting measure: inference on the intensity process**. Chemometrics and Intelligent Laboratory, 103, 116-121, 2010b.
- [27] Bouzas PR, Ruiz-Fuentes N, Ruiz-Castro JE. **Forecasting a compound Cox process by means of PCP**. Proceedings in Computational Statistics, Physica-Verlag, Berlin, 2010c.
- [28] Bouzas PR. **Proceso de Cox. Del proceso de Poisson a generalizaciones del proceso de Cox**. Editorial académica española, 2011.
- [29] Bouzas PR, Aguilera AM, Ruiz-Fuentes N. **Functional estimation of the random rate of a Cox process**. Methodology and Computing in Applied Probability, 14, 57-69, 2012.
- [30] Bouzas PR, Ruiz-Fuentes N, Ruiz-Castro JE. **Forecasting compound Cox processes**. Invited speaker in the 7th International Conference of

- the ERCIM WG on Computational and Methodological Statistics (ERCIM 2014), University of Pisa, Italy, 2014.
- [31] Bouzas PR, Ruiz-Fuentes N. **A Review on Functional Data Analysis for Cox processes**. BEIO (Boletín de Estadística e Investigación Operativa), 31 ,3, 215 - 230, 2015.
- [32] Bouzas PR, Ruiz-Fuentes N, Montes-Gijón C. **Phase type process**. Invited speaker in the 10th International Conference of the ERCIM WG on Computational and Methodological Statistics (ERCIM 2017), University of London, United Kingdom, 2017.
- [33] Bouzas PR, Ruiz-Fuentes N, Montes-Gijón C, Ruiz-Castro JE. **Forecasting counting and time statistics of compound Cox processes: a focus on intensity phase type process, deletions and simultaneous events**. Statistical Paper. <https://doi.org/10.1007/s00362-019-01092-0>, 2019.
- [34] Brémaud P. **Point processes and queues: martingale dynamics**. Springer, New York, 1981.
- [35] Brix A. **Spatio-temporal modeling of weeds by shot-noise G Cox processes**. Biometrical Journal, 44, 83-99, 2002.
- [36] Chen F, Hall P. **Inference for a non-stationary Self-exciting point process with an application in ultra-high frequency financial data modeling**. Journal of Applied Probability, 50, 4, 1006-1024, 2013.
- [37] Chertok AV, Korolev VY, Korchagin AY. **Modeling high-frequency non-homogeneous order flows by compound Cox processes**. Journal of Mathematical Sciences, 214, 1, 44-68, 2016.

- [38] Chiang CT, Wang MC, Huang CY. **Kernel estimation of rate function for recurrent event data**. Scandinavian Journal of Statistics, 32, 77-91, 2005.
- [39] Cook RJ, Lawless JF. **The statistical analysis of recurrent events**. Springer-Verlag, New York, 2007.
- [40] Cox DR. **Some statistical methods connected with series of events**. Journal of the Royal Statistical Society: Series B, 17, 129-164, 1955.
- [41] Daley DJ, Vere-Jones D. **An introduction to the theory of point processes**. Springer-Verlag, New York, 1988.
- [42] Daley DJ, Vere-Jones D, **An introduction to theory of point processes**. Second Edition, Springer-Verlag, New York, 1993.
- [43] Davenport Jr WB, Root WL. **An introduction to the theory of random signals and noise**. IEEE Communication Society, Wiley-Interscience, New York, 1958.
- [44] Dayar T, **On moments of discrete Phase-type distributions**. European Workshop on Performance Engineering, International Workshop on Web Services and Formal Methods, Formal Techniques for Computer Systems and Business Processes, Lecture Notes in Computer Science, 3670, 51-63, Springer-Verlag, Berlin, Heidelberg, 2005.
- [45] Diggle PJ, Heagerty P, Liang K, Zeger S, **Analysis of longitudinal data**. Second Edition, Oxford Statistical Science Series, Oxford University Press, Oxford, 2012.

-
- [46] Dousse O, Baccelli F, Thiran P, **Impacto of interferences on connectivity in ad hoc networks**. IEEE/ACM Transactions on Networking, 13, 425-436, 2005.
- [47] Economou A, **On the control of a compound immigration process through total catastrophes**. European Journal of Operational Research- Elsevier, 147, 522-529, 2003.
- [48] Erlang AK, **Sandsynlighedsregning og telefonsamtaler (The theory of probabilities and telephone conversations)**. Nyt Tidsskrift for Matematik, Afdeling B, 20, 33-39, 1909.
- [49] Faenza L, Hainzl S, Scherbaum F, **Statistical analysis of the Central-Europe seismicity**. Tectonophysics-Elsevier, 53, 4322-4331, 2008.
- [50] Fackrell M, **Modelling healthcare systems with phase-type distributions**. Health Care Management Science-Springer Science + Business Media, LLC, 12, 11-26, 2008.
- [51] Fernández-Alcalá R, Navarro-Moreno J, Ruiz-Molina JC, **Statistical inference for doubly stochastic multichannel Poisson processes: a PCA approach**. Computational Statistics and Data Analysis, 53, 12, 4322-4331, 2009.
- [52] Fernández-Alcalá R, Navarro-Moreno J, Ruiz-Molina JC, **On the smoothing estimation problem for the intensity of a DSMPP**. Methodology and Computing in Applied Probability, 14, 1, 5-16, 2012.
- [53] Fritsch FN, Carlson RE, **Monotone piecewise cubic interpolation**. SIAM Journal on Numerical Analysis, 7, 238-246, 1980.

- [54] García R, Ghysels E, Renault E, **The econometrics of option pricing**. CIRANO Scientific Series. 2004.
- [55] Gastón M, León T, Mallor F, **Functional data analysis for non homogeneous Poisson processes**. Proceedings of the 2008 Winter Simulation Conference, Miami. 2008.
- [56] Genaro AD, Simonis A, **Estimating doubly stochastic poisson process with affine intensities by kalman filter**. Statistical Papers 56, 723-748, 2015.
- [57] Gilardoni GL, Colosimo EA, **On the superposition of overlapping Poisson processes and nonparametric estimation of their intensity function**. Journal of Statistical Planning and Inference, 141, 3075-3083, 2011.
- [58] Gospodinov D, Rotondi R, **Exploratory analysis of marked Poisson processes applied to Balkan earthquake sequences**. Journal of the Balkan Geophysical Society, 4, 61-68, 2001.
- [59] Grandell J, **On stochastic processes generated by a stochastic intensity function**. Skandinavisk Aktuarietidskrift, 3, 4, 204-240. 1971.
- [60] Greenberg D, Houweling A, Kerr J, **Population imaging of ongoing neuronal activity in the visual cortex of awake rats**. Nature Neuroscience, 11, 749-751, 2008.
- [61] Greenwood M, Yule GU, **An enquiry into the nature of frequency distributions of multiple happenings, with particular reference to the occurrence of multiple attacks of disease or repeated accidents**. Journal of the Royal Statistical Society, 83, 255-279, 1920.

-
- [62] Haas SM, Shapiro JH, **Capacity of the multiple-input, multiple-output Poisson channel**. Stochastic Theory and Control, 280, 155-168, 2002.
- [63] Hall P, Müller H, Yao F, **Modelling sparse generalized longitudinal observations with latent Gaussian processes**. Journal of the Royal Statistical Society: Series B, 70, 703-723, 2008.
- [64] Hätler G, **The nonhomogeneous Poisson process-A model for the reliability of complex repairable systems**. Microelectronics & Reliability, 9, 3, 381-386, 1989.
- [65] Hawkes AG, **Spectra of some self-exciting an mutually exciting point processes**. Biometrika, 58, 1, 83-90, 1971.
- [66] He QM, **Fundamental of matrix-analitics methods**. Springer Science + Business Media, New York, 2014.
- [67] Ho LP, Stoyan D. **Modelling marked point patterns by intensity-marked Cox process**. Statistics and Probability Letters, 78, 1194-1199, 2008.
- [68] Hutchison K, Willemain TR, **Nonparametric methods for forecasting the Cox process**. Proceedings of the 2013 Industrial and Systems Engineering Research Conference. 2013.
- [69] Hautsch N, **Econometrics of financial high-frequency data**. Springer Heidelberg Dordrecht London New York, 284-289, 2012.
- [70] Jiménez F, **Contribución al estudio de procesos poissonianos de recuento con intensidad aleatoria**. Tesis, Universidad de Granada, 1991.

-
- [71] Johnson NL, Kotz S, **Continuous univariate distributions-I, distributions in statistics**. Wiley, New York, 1970.
- [72] Kloeden PE, Platen E, **Numerical solution of stochastic differential equations**. Springer, Berlin, 1999.
- [73] Kolmogorov AN, **La transformation de Laplace dans les espaces linaires**. Comptes Rendus de L'Académie des Sciences, 200, 1717-1718, 1935.
- [74] Komance A, **Optimal detection of marked and filtered Poisson process in additive white gaussian noise**. 45th Annual Allerton Conference, Illinois, 2007.
- [75] Kopperschmidt K, Stute W, **The statistical analysis of self-exciting point processes**. *Statistica Sinica*, 23, 1273-1298, 2013.
- [76] Last G, Brandt A, **Marked point processes on the real line**. (The dynamic approach). Springer-Verlag, New York, 1995.
- [77] Lefebvre M, Bensalma F, **Modeling and forecasting river flows by means of filtered Poisson processes**. *Applied Mathematical Modelling*, 39, 230-243, 2015.
- [78] Lèveille G, Garrido J, **Moments of compound renewal sums with discounted claims**. *Insurance: Mathematics and Economics*, 28, 217-231, 2001.
- [79] Lewis E, Mohler G, Brantingham PJ, Bertozzi AL, **Self-exciting point process models of civilian deaths in Iraq**. *Security Journal* 25, 3, 244-264, 2012.

-
- [80] Lin XS, Pavlova KP, **The compound Poisson risk model with a threshold dividend**. Insurance: Mathematics and Economics, 38, 1, 57-80, 2006.
- [81] Lin Z, Huang Y, **Nonperiodic preventive maintenance for repairable systems**. Naval Research Logistics, 57, 615-625, 2010.
- [82] Maier R, O’Cinneide CA, **A closure characterization of phase-type distribution**. Journal of Applied Probability, 29, 1, 92-103, 1992.
- [83] Minc H, **Non-negative matrices**. Wiley, New York, 1988.
- [84] Myllymaki M, Pettinen A, **Bayesian inference for gaussian excursion set generated Cox processes with set-marking**. Journal of Statistics and Computing, 20, 315-320, 2010.
- [85] Mohler GO, Short MB, Brantingham PJ, Schoenberg FG, Tita GE, **Self-exciting point process modeling of crime**. Journal of the American Statistical Association 106, 493, 100-108, 2011.
- [86] Møller J, Waagepeterse RP, **Statistical inference and simulation for spatial point processes**. Chapman and Hall, New York, 2004.
- [87] Montes-Gijón C, Bouzas PR, Ruiz-Fuentes N, **Forecasting compound Cox processes**. Invited speaker in the 9th International Conference of the ERCIM WG on Computational and Methodological Statistics (ERCIM 2016), University of Seville, Spain, 2016.
- [88] Moyal JE, **The general theory of stochastic population processes**. Acta Mathematica, 108, 1-31, 1962.

-
- [89] Neuts MF, **Probability distributions of phase type**. Liber Amicorum Prof. Emeritus H. Florin, Department of Mathematics, University of Louvain, Belgium, 173-206, 1975.
- [90] Neuts MF, **Matrix-geometric solutions in stochastic models: an algorithmic approach**. Dover Publications Inc, New York, 1981.
- [91] O’Cinneide CA, **Characterization of phase-type distributions**. Stochastic Models, 6, 1-57, 1990a.
- [92] Ogata Y, **Space-time point-process models for earthquake occurrences**. Annals of the Institute of Statistical Mathematics, 50, 379-402, 1998.
- [93] Ogata Y, Katsura K, Zhuang J, **TIMSAC84: statistical analysis of series of events (TIMSAC84-SASE) Version 2**. Computer Science Monographs, 32, 2006.
- [94] Park C, Padgett W, **Accelerated degradation models for failure based on geometric brownian motion and gamma processes**. Lifetime Data Analysis, 11, 511-527, 2005.
- [95] Peña EA, Strawderman RL, Hollander M, **Nonparametric estimation with recurrent event data**. Journal of the American Statistical Association, 96, 1299-1315, 2001.
- [96] Pham H, **Handbook of reliability engineering**. Springer-Verlag London Limited, 2003.
- [97] Poisson, SD. **Rechercher sur la Probabilité des Jugements en Matière Criminelle et en Matière Civile, Précédées des Règles Gènèrales du Calcul des Probabilitès**. Bachelier, Paris, 1837.

-
- [98] Reineck P, Bodrog L, Danilkina A, **Phase type distributions**. K. Wolter et al. (eds.), Resilience Assessment and Evaluation of Computing Systems, Springer-Verlag Berlin Heidelberg, 2012.
- [99] Renshaw E, Comas C, Mateu J, **Analysis of forest thinning strategies through the development of space-time growth-interaction simulation models**. Stochastic Environmental Research and Risk Assessment, 23,275-288, 2009.
- [100] Rolski T, Schmidli H, Smhmidt V, Teugels J, **Stochastic processes for insurance and finance**. Wiley, New York, 2001.
- [101] Ruiz-Castro JE, **Complex multi-state systems modelled through marked markovian arrival processes**. European Journal of Operational Research, 252,3, 852-865, 2016.
- [102] Ruiz-Castro JE, **Markov counting and reward processes for analyzing the performance of a complex system subjecto to random inspections**. Reliability Engineering and System Safety, 145,155-168, 2016.
- [103] Ruiz-Fuentes N, Bouzas PR, Atsalakis G, **Turning points modelled as a Cox process**. Book of abstracts of the 8th International Conference on Computational Management Science (CMS2011), Neuchâtel, Switzerland, 2011.
- [104] Ruiz-Fuentes N, Bouzas PR, **Characterizing self-exciting processes: simulations driven by phase-type distributions**. Invited speaker in the 6th International Conference of the ERCIM WG on Computational and Methodological Statistics (ERCIM 2013), University of London, United Kingdom, 2013.

-
- [105] Ruiz-Fuentes N, Bouzas PR, Montes-Gijón C, **Self-exciting process with phase-type intensity: simulation and estimation**. Invited speaker in the 7th International Conference of the ERCIM WG on Computational and Methodological Statistics (ERCIM 2014), University of Pisa, Italy, 2014.
- [106] Ruiz-Fuentes N, Bouzas PR, **Goodness-of-fit test for Cox processes and generalizations**. Invited speaker in the 9th International Conference of the ERCIM WG on Computational and Methodological Statistics (ERCIM 2016), University of Seville, Spain, 2016.
- [107] Ruiz-Fuentes N, Bouzas PR, Montes-Gijón C, **Extreme events of meteorological parameters modeled by Cox processes**. XVI Conferencia Española de Biometría. CEB2017. Sevilla, 2017.
- [108] Ruiz-Fuentes N, Bouzas PR, Montes-Gijón C, **Compound Cox processes applied to extreme meteorological events**. Invited speaker in the 10th International Conference of the ERCIM WG on Computational and Methodological Statistics (ERCIM 2017), University of London, United Kingdom, 2017.
- [109] Russell JR, Engle RF, **Analysis of high-frequency data**. Handbook of financial econometrics: tools and techniques, Elsevier, 383-426, 2010.
- [110] Seidel H, **Über die Probabilitäten solcher Ereignisse welche nur seiten vorkommen, obgleich sie unbeschränkt oft möglich sind**. Sitzungsberichte der Mathematisch-Physikalischen Classe der K.B., Akademie der Wissenschaften zu München, 6, 44-50, 1876.

-
- [111] Sepehrifar M, Yarahmadian S, **Decreasing renewal dichotomous Markov noise shock model with hypothesis testing applications**. Statistical Papers 58, 1115-1124, 2017.
- [112] Serfozo RF, **Conditional Poisson processes**. Journal of Applied Probability, 9, 288-302, 1972.
- [113] Shi DH, Guo J, Liu L, **SPH-distributions and rectangle-iterative algorithm**. Matrix-analytic methods in stochastic models. Marcel Dekker, New York, 207-224, 1996.
- [114] Si S, **Random irreversible phenomena: entropy insubordination**. Chaos, Solitons & Fractals, 12, 2873-2876, 2001.
- [115] Singh S, Tripathi YM, **Estimating the parameters of an inverse Weibull distribution under progressive type-i interval censoring**. Statistical Papers 12, 2873-2876, 2018.
- [116] Snyder DL, **Random point processes**. Wiley, New York, 2, 1975.
- [117] Snyder DL, Miller MI, **Random point processes in time and space**. Second edition, Springer-Verlag, New York, 1991.
- [118] Spencer QH, Jeffs BD, Jensen MA, Swindlehurst AL, **Modeling the statistical time and angle of arrival characteristics of an indoor multipath channel**. IEEE Journal on Selected Areas in Communications, 18, 347-360, 2000.
- [119] Spreij P, **Self-exciting counting process systems with finite state space**. Stochastic Processes and their Applications 34, 275-395, North-Holland, 1990.

-
- [120] Srinivasan SK, Vasudevan R, **Introduction to random differential equations and their applications**. Elsevier, New York, 1971.
- [121] Tank F, Eryilmaz S, **The distributions of sum, minima and maxima of generalized geometric random variables**. Statistical Papers 56, 1191-1203, 2015.
- [122] Valderrama MJ, Jiménez F, Gutiérrez M, Martínez-Almécija A, **Estimation and filtering on a doubly stochastic Poisson process**. Applied Stochastic Models and Data Analysis, 11, 13-24, 1995.
- [123] Wu S, Müller H, Zhang Z, **Functional data analysis for point processes with rare events**. Statistica Sinica, 23, 1-23, 2013.
- [124] Wang R, **A mixture and self-exciting model for software reliability**. Statistics & Probability Letters, 72,187-194, 2005.
- [125] Zhang T, Kou SC, **Nonparametric inference of doubly stochastic Poisson process data via the Kernel method**. Annals of Applied Statistics, 4, 1913-1941, 2010.
- [126] Zhao M, Xie , **On maximum likelihood estimation for a general non-homogeneous Poisson process**. Scandinavian Journal of Statistics, 23, 597-607, 1996.
- [127] Zhao X, Balakrishnan N, Sun J, **Nonparametric inference based on panel count data**. Test, 20, 1, 1-42, 2011.